CONDIÇÕES DE OTIMALIDADE PARA OS PROBLEMAS FINITO- DIMENSIONAL E DE TEMPO CONTÍNUO

LUCELINA BATISTA DOS SANTOS

Prof. Dr. Marko Antonio Rojas Medar
Orientador

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para a obtenção do título de Mestre em Matemática Aplicada.

IMECC- UNICAMP
ABRIL-2000

CONDIÇÕES DE OTIMALIDADE PARA OS PROBLEMAS FINITO- DIMENSIONAL E DE TEMPO CONTÍNUO

LUCELINA BATISTA DOS SANTOS

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Lucelina Batista dos Santos e aprovada pela Comissão Julgadora.

Banca Examinadora

Prof. Dr. Marko Antonio Rojas Medar (orientador)
Prof. Dr. Antonio Carlos Moretti
Prof. Dr. Adilson José Vieira Brandão

Campinas, 07 de Abril de 2000.

Prof. Dr. Marko Antonio Rojas Medar
(orientador)

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para obtenção do Título de Mestre em Matemática Aplicada.

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP

Santos, Lucelina Batista dos

Sa59c Condições de otimalidade para os problemas finito dimensional e de tempo contínuo / Lucelina Batista dos Santos -- Campinas, [S.P. :s.n.], 2000.

Orientador : Marko Antonio Rojas Medar

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Otimização matemática. 2. Programação não-linear. 3. Otimização não diferenciável. I. Rojas Medar, Marko Antonio. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Dissertação de Mestrado defendida em 07 de abril de 2000 e aprovada

pela Banca Examinadora composta pelos Profs. Drs.

____

Prof (a). Dr (a). MARKO ANTONIO ROJAS MEDAR

Antonio Carlos Moretti

Prof (a). Dr (a). ANTONIO CARLOS MORETTI

____

Prof (a). Dr (a). ADILSON JOSÉ VIEIRA BRANDÃO

Dedicatória

Este trabalho é dedicado aos meus pais, Sr. Luiz e D. Celina.

Agradecimentos

- À FAPESP pelo suporte financeiro e pelas críticas recebidas durante a execução deste trabalho;

- Aos membros da Banca Examinadora, pelas sugestões que em muito contribuíram para a redação final desta dissertação;

- Aos funcionários do IMECC; em particular à Fátima da Secretaria da Matemática Aplicada e à Cidinha da Secretaria de Pós Graduação pelo auxílio prestado de forma competente e amável;

- Aos amigos e colegas, pela "força". Em particular, agradeço ao Yurilev e ao meu irmão Júnior pela ajuda na digitação e à amiga Patrícia pelo "apoio incondicional";

- Ao Prof. Marko pela proposta de trabalho, pela orientação segura e por sua amizade.

Resumo

Neste trabalho, um teorema de alternativa do tipo Gordan é utilizado no estudo de condições necessárias de otimalidade para o problema clássico de programação não linear (finito- dimensional). Condições suficientes são obtidas através de uma noção de convexidade generalizada (chamada invexidade). Além disso, sem hipóteses de convexidade (generalizada ou não) são obtidas condições suficientes de otimalidade via método de deformação.

Resultados análogos são válidos para o problema de tempo contínuo (exceto o método de deformação).

Abstract

In this work, an alternative theorem (Gordan's type) is used to investigate necessary conditions of optimality for the classic nonlinear programming problem (finite dimensional). Sufficient conditions are obteined via a notion of generalized convexity (called invexity) . Moreover, without hipothesis of convexity (generalized or no) are obtained sufficient conditions via deformation method.

Analogous results are valid for the continous time problem (except for the method of deformation).

Índice

Capítulo 1

Introdução e preliminares técnicos

1.1 Introdução

Neste trabalho estudamos, quanto a condições de otimalidade, dois problemas: o Problema de Programação Não Linear (em dimensão finita) e também o chamado Problema de Tempo Contínuo, os quais formulamos a seguir.

Iniciamos com a formulação do problema finito- dimensional:

$$\begin{aligned} &\text{minimizar } f_0(x) \\ &\text{sujeito a } f_i(x) \leq 0, \;\; i = 1, ..., m \end{aligned} \tag{P}$$

onde $f_0, f_i : X \rightarrow \mathbb{R}$ são funções definidas em um subconjunto X de $\mathbb{R}^n$.

Necessitaremos das seguintes definições:

Definição 1.1 :

1. *$x^* \in X$ é um **ponto factível** para (P) se $f_i(x^*) \leq 0$ para todo $i = 1, ..., m$ e denotaremos o conjunto dos pontos factíveis por $\mathbb{F}$.*

2. *$x^* \in X$ é uma **solução local** (ou mínimo local) para (P) se é factível e se existe uma vizinhança V de x^* tal que $f_0(x^*) \leq f_0(x)$ para todo $x \in V \cap \mathbb{F}$.*

3. *$x^* \in X$ é uma **solução global** de (P) se é factível e se $f_0(x^*) \leq f_0(x)$ para todo $x \in \mathbb{F}$.*

Se X e Y forem espaços de Banach, o conceito de continuidade de funções pode ser generalizado como segue:

Definição 1.3 *Sejam X e Y espaços de Banach, $\Gamma : X \rightrightarrows Y$ uma multifunção e $x' \in X$. Dizemos que Γ é* ***semicontínua superior*** *em x' se para cada $\varepsilon > 0$, existe um $\delta > 0$ tal que*

$$\Gamma(x) \subset \Gamma(x') + \varepsilon B_Y \text{ para todo } x \in \{x'\} + \delta B_X.$$

(onde B_X e B_Y são as bolas abertas unitárias dos espaços X e Y)

Quando não houver perigo de confusão, denotaremos, simplesmente, por B a bola unitária aberta do espaço considerado.

Seja dado D um subconjunto não vazio de um espaço de Banach B. Definimos o **funcional suporte** de D como sendo $\sigma_D : B^* \to \mathbb{R} \cup \{\infty\}$, onde $\sigma_D(\xi) = \sup\{< \xi, x >: x \in D\}$. (Onde $< \cdot, \cdot >$ denota a relação de dualidade canônica em $B^* \times B$).

Pode-se provar que se C, D são subconjuntos não vazios, fechados e convexos de B e se μ, λ são não negativos, então valem as seguintes propriedades (Hörmander):

1. $C \subset D$ se e somente se $\sigma_C(\xi) \leq \sigma_D(\xi)$ para todo $\xi \in B^*$;

2. $\mu\sigma_C(\xi) + \lambda\sigma_D(\xi) = \sigma_{(\mu C + \lambda D)}(\xi)$ para todo $\xi \in B^*$.

Outra noção de que necessitaremos é a de mensurabilidade de uma multifunção.

Definição 1.4 *Seja $G : [0,T] \rightrightarrows \mathbb{R}^n$ uma multifunção. G é* ***mensurável*** *se para cada subconjunto fechado C de $\mathbb{R}^n$ o conjunto $\Gamma(C) = \{x \in [0,T] : \Gamma(x) \cap C \neq \phi\}$ é Lebesgue mensurável.*

Observamos que a Definição 1.4, de fato, generaliza a noção clássica de mensurabilidade de funções.

Em contexto clássico, uma função $f : [0,T] \to \mathbb{R}^n$ é mensurável se para cada subconjunto fechado C de $\mathbb{R}^n$, temos que $f^{-1}(C)$ é Lebesgue mensurável.

Façamos $\widetilde{f}(x) = \{f(x)\}$ e consideremos a multunção $\widetilde{f} : [0,T] \rightrightarrows \mathbb{R}^n$.

Neste caso:

2. $x^* \in X$ é uma ***solução local*** *(ou mínimo local) para (CNP) se é factível e se existe uma vizinhança V de x^* tal que $\phi(x^*) \leq \phi(x)$ para todo $x \in V \cap \mathbf{F}$.*

3. $x^* \in X$ é uma ***solução global*** *de (CNP) se é factível e se $\phi(x^*) \leq \phi(x)$ para todo $x \in \mathbf{F}$.*

Veremos que resultados análogos aos obtidos para o problema finito- dimensional podem ser obtidos para o problema (CNP).

Em [7], Bobylev obtém, através do método de deformação, condições suficientes de otimalidade para o problema finito- dimensional, sem nenhuma hipótese de convexidade (generalizada ou não).

Nossa tentativa foi adaptar esta argumentação ao problema (CNP) e, de forma análoga, obtermos condições suficientes de otimalidade para o problema (CNP) sem hipóteses adicionais de convexidade. Não logramos atingir este resultado e os motivos disto são expostos no Capítulo 7 (Parte II) deste trabalho

1.2 Preliminares técnicos

Nesta seção apresentamos alguns resultados de análise não diferenciável e de convexidades generalizadas, os quais serão utilizados nos capítulos posteriores deste trabalho. São fatos conhecidos e que, por isto, não serão demonstrados aqui . Estes podem ser encontrados, respectiamente, no livro de Aubin e Frankowska [2] e no (excelente!) livro de Clarke [3].

1.2.1 Alguns resultados de Análise Multívoca

Em nossa discussão, necessitaremos de algumas definições de Análise Multívoca, as quais listaremos a seguir. Iniciamos com o conceito de multifunção:

Sejam X e Y dois subconjuntos não vazios. Por uma **multifunção** de X em Y entendemos uma função Γ definida em X e com valores no conjunto das partes de Y, e a denotaremos por $\Gamma : X \rightrightarrows Y$.

Um exemplo de multifunção é o gradiente generalizado de Clarke, o qual definimos na próxima subseção e que será de grande importância neste trabalho, como se verá nos capítulos seguintes.

4. *Para cada $x \in \mathbb{F}$ definimos o conjunto $I(x) = \{i = 1, ..., m : f_i(x) = 0\}$ chamado o* ***conjunto das restrições ativas*** *em x.*

Veremos que sob certas hipóteses de diferenciabilidade sobre os funcionais do problema (P), é possível encontrar condições necessárias de otimalidade sob a forma de uma regra de multiplicadores. Veremos que se os funcionais do problema possuem algum tipo de convexidade generalizada, essas condições se tornarão, também, suficientes.

Formulamos a seguir o Problema de Tempo Contínuo:

$$\left.\begin{array}{l} \text{minimizar } \phi(x) = \int_0^T f(t, x(t))dt \\ \text{sujeito a: } g_i(t, x(t)) \leq 0 \text{ q.t.p. em } [0, T] \\ i \in I = \{1, ..., m\} \\ x \in X \end{array}\right\} \qquad \text{(CNP)}$$

onde X é um subconjunto não vazio, aberto e convexo do espaço $L_n^\infty[0, T]$ das funções definidas no intervalo $[0, T]$ a valores em $\mathbb{R}^n$, Lebesgue- mensuráveis, essencialmente limitadas. Este espaço, munido da norma $||\cdot||_\infty$ é um espaço de Banach, com esta norma definida por

$$||x||_\infty = \max_{j=1,...,n} ess\sup\{|x_j(t)|;\ t \in [0, T]\},\ x \in L_n^\infty[0, T]$$

Admitiremos ainda, que $f(t, x(t)) = \Gamma(x)(t)$ e $g(t, x(t)) = (g_1(t, x(t)), ..., g_m(t, x(t))) = \gamma(x)(t)$ onde $\gamma : X \to \Lambda_1^m[0, T]$ e $\Gamma : X \to \Lambda_1^1[0, T]$; o conjunto $\Lambda_1^m[0, T]$ é constituído das funções definidas em $[0, T]$ com valores em $\mathbb{R}^m$, Lebesgue mensuráveis, essencialmente limitadas e consideraremos o espaço $\Lambda_1^m[0, T]$ munido da norma $||\cdot||_1$, definida por

$$||y||_1 = \max_{j=1,...,m} \int_0^T |y_j(t)|dt,\ x \in \Lambda_1^m[0, T].$$

De maneira análoga ao problema (P), definimos:

Definição 1.2 *Seja o problema (CNP) definido acima.*

1. *$x^* \in X$ é uma* ***solução factível*** *para (CNP) se $g_i(t, x(t)) \leq 0$ q.t.p. em $[0, T]$, para todo $i \in I = \{1, ..., m\}$ e denotaremos o conjunto das soluções factíveis de (CNP) por $\mathbb{F}$.*

$$
\begin{aligned}
\Gamma(C) &= \{x \in [0,T] : \widetilde{f}(x) \cap C \neq \phi\} \\
&= \{x \in [0,T] : \{f(x)\} \cap C \neq \phi\} \\
&= \{x \in [0,T] : f(x) \in C\} \\
&= f^{-1}(C)
\end{aligned}
$$

e, portanto, uma função f é mensurável se e somente se a multifunção associada $\widetilde{f}$ é mensurável no sentido da Definição 1.4.

A noção de integral de Lebesgue é generalizada para o contexto multívoco da seguinte maneira:

Definição 1.5 *Sejam* $G : [0,T] \rightrightarrows \mathbb{R}^n$ *uma multifunção e o conjunto* $S^1([0,T]) = \{f \in L_1^n[0,T] : f(t) \in G(t)$ *q.t.p. em* $[0,T]\}$. *A **integral** (de Aumann) de* G *é definida como sendo o conjunto*

$$\int_0^T G(t)dt = \{\int_0^T f(t)dt : f \in S^1([0,T])\}.$$

Dizemos que multifunção $G : [0,T] \rightrightarrows \mathbb{R}^n$ é **integravelmente limitada** se G é mensurável e se existe uma função $z : [0,T] \to \mathbb{R}_+$ integrável tal que $||G(t)|| \leq z(t)$ q.t.p. em $[0,T]$.

Pode-se provar (v. Aubin e Frankowska, [2]) que se a multifunção G é integravelmente limitada e assume valores compactos em $\mathbb{R}^n$, então,

$$\int_0^T \sigma_{G(t)}(v)dt = \sigma_{\int_0^T G(t)dt}(v), \text{ para todo } v \in \mathbb{R}^n.$$

1.2.2 Ferramentas de Análise não diferenciável

Ao longo desta subseção, assumiremos $(E, ||\cdot||)$ um espaço de Banach; denotaremos seu dual topológico por E^* e o adimitiremos munido da norma

$$||\xi||_* = \sup\{< \xi, v >: v \in E, ||v|| \leq 1\}.$$

Assumiremos, ainda, que Ω é um subconjunto não vazio aberto de E.

Seja $f : \Omega \to \mathbb{R}$ um funcional. Dizemos que f é **Lipschitz próximo de** $x \in \Omega$ se existem um $\delta > 0$ e um $k = k(x, \delta)$ tal que

$$|f(x_1) - f(x_2)| \leq k||x_1 - x_2||$$

para quaisquer $x_1, x_2 \in \Omega \cap (\{x\} + \delta B)$ (onde B é a bola unitária em E).

Além disso, se f é Lipschitz próximo de x, para todo $x \in \Omega$, dizemos que f é **localmente Lipschitz** em Ω.

Definição 1.6 *Sejam $x \in \Omega$ e $v \in E$. Definimos a* ***derivada direcional*** *(de Clarke) de f em x na direção v como sendo*

$$f^0(x; v) = \limsup_{\substack{y \to x \\ \lambda \downarrow 0}} \frac{1}{\lambda}[f(y + \lambda v) - f(y)]$$

(onde $y \in \Omega$, $\lambda \in (0, +\infty)$).

Definição 1.7 *Definimos o* ***gradiente generalizado*** *(de Clarke) de f em $x \in \Omega$ como sendo o conjunto*

$$\partial f(x) = \{\xi \in E^* : f^0(x; v) \geq < \xi, v >, \forall v \in E\}.$$

O teorema seguinte nos dá algumas propriedades da derivada direcional e do gradiente generalizados:

Teorema 1.8 *Seja $f : \Omega \to \mathbb{R}$ um funcional localmente lipschitz com constante k. Então as seguintes afirmações são verdadeiras:*

1. *A função $v \mapsto f^0(x; v)$ é finita, sublinear e satisfaz $|f^0(x; v)| \leq k||v||$;*
2. *Para cada $x \in \Omega$, $\partial f(x)$ é um subconjunto não vazio, convexo e w^*-compacto em E^*; além disso, $||\xi||_* \leq k$, para todo $\xi \in \partial f(x)$;*

3. *Para cada* $v \in E$, *temos* $f^0(x;v) = \max\{<\xi, v>: \xi \in \partial f(x)\}$;

4. $\xi \in \partial f(x)$ *se e somente se* $f^0(x;v) \geq <\xi, v>$ *para todo* $v \in E$.

O resultado anterior é exemplo de um fato mais geral: conjuntos convexos e fechados são caracterizados por seus funcionais suporte. Deste modo, o teorema anterior afirma que $f^0(x;v) = \sigma_{\partial f(x)}(v)$ e, portanto, conhecer $\partial f(x)$ equivale a conhecer $f^0(x;\cdot)$.

Outro conceito que nos será útil é o de regularidade o qual definimos a seguir:

Definição 1.9 *O funcional* $f : \Omega \to \mathbb{R}$ *é* ***regular*** *(ou Clarke regular) em* $x \in \Omega$ *se*

1. *Para todo* $v \in E$, *existe a derivada direcional (usual) de* f *em* x *na direção* v, $f'(x;v)$;

2. *Para cada* $v \in E$, $f'(x;v) = f^0(x;v)$.

Além disso, se f *é regular em todo* $x \in \Omega$ *então* f *é* ***regular*** *(em* Ω*).*

Pode-se demonstrar:

(a) Se f_i são regulares em $x \in \Omega$, então $\partial(\sum_{i=1}^{n} f_i)(x) = \sum_{i=1}^{n} \partial f_i(x)$. (A inclusão $\subset$ vale sempre.)

(b) Se f é lipschitz próximo de $x \in \Omega$ é convexa, então f é regular.

(c) Se f é continuamente diferenciável em $x \in \Omega$, então $\partial f(x) = \{f'(x)\}$ (onde f' é a derivada usual). [Isto mostra que esta é uma boa generalização de diferenciabilidade.]

Agora, observamos que se C é um subconjunto não vazio de E, então a **função distância** $d_C : E \to \mathbb{R}$ definida por $d_C(x) = \inf\{||x - c|| : c \in C\}$ não é diferenciável, mas é globalmente Lipschitz, com constante de Lipschitz igual a 1. (v. [13], p. 50). Feita esta consideração, definimos

Definição 1.10 *Sejam* C *um subconjunto não vazio de* E *e seja* $x \in C$. *Dizemos que* $v \in E$ *é um* ***vetor tangente*** *a* C *em* x *se* $d^0(x;v) = 0$. *Denotaremos* $T_C(x)$ *o conjunto dos vetores tangentes a* C *em* x *e o chamaremos* ***cone tangente*** *(de Clarke). Definimos, ainda, o cone normal a* C *em* x *como sendo o conjunto*

$$N_C(x) = \{\xi \in E^* : < \xi, v > \leq 0, \forall v \in T_C(x)\}.$$

Pode-se provar que para cada $x \in C$, $T_C(x)$ é um cone convexo e fechado em E e que $N_C(x)$ é um cone convexo e w*-fechado em E^*.

O teorema seguinte (que se encontra demonstrado em [13], p. 63), nos dá uma boa caracterização de $\partial f(x)$ no caso $E = \mathbb{R}^n$:

Teorema 1.11 *Sejam Ω um subconjunto não vazio de $\mathbb{R}^n$, $f : \Omega \to \mathbb{R}$, S um subconjunto de medida nula em $\mathbb{R}^n$ e Ω_f o conjunto dos pontos onde f não é diferenciável. Então:*

$$\partial f(x) = \mathrm{co}\{\lim \nabla f(x_i) : x_i \to x, x_i \notin S, x_i \notin \Omega_f\}.$$

A minimalidade de um funcional em um conjunto pode ser expressa pela seguinte condição estacionária:

Teorema 1.12 *(condição necessária de otimalidade) Sejam $f : \Omega \to \mathbb{R}$ localmente Lipschitz, x^* um mínimo de f em C. Então*

$$0 \in \partial f(x^*) + N_C(x^*) \qquad (*)$$

(e um ponto $x^0 \in C$ que satisfaz (*) é um ***ponto estacionário*** *(generalizado) ou* ***extremo de*** *f)*

No caso em que $C = \Omega = E$, temos $T_C(x^*) = E$, e, portanto, $N_C(x^*) = \{0\}$, e (*) toma a forma mais conhecida, $0 \in \partial f(x^*)$.

1.2.3 Convexidades Generalizadas

A convexidade em Programação Matemática é muito importante, pois garante critérios de suficiência otimal. Nesta seção definimos alguns tipos de convexidade generalizada e a relacionaremos com a noção clássica de convexidade.

Nas definições seguintes assumiremos que S é um subconjunto convexo, não vazio de um espaço normado.

Recordemos que uma função $f : S \to \mathbb{R}$ é **convexa** se para quaisquer $x_1, x_2 \in S$ e para todo $\lambda \in [0, 1]$ temos

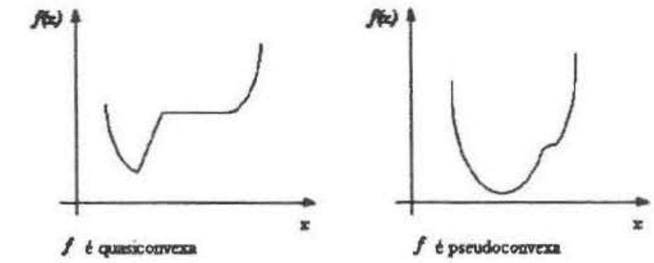


Figure 1-1:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2).$$

É fato que, se a função f é convexa, então, para cada $\alpha \in \mathbb{R}$, o conjunto $\Lambda_\alpha = \{x \in S : f(x) \leq \alpha\}$ é convexo. Isto motiva a seguinte

Definição 1.13 *Dizemos que a função $f : S \to \mathbb{R}$ é* ***quasiconvexa*** *se para cada $\alpha \in \mathbb{R}$, o conjunto $\Lambda_\alpha = \{x \in S : f(x) \leq \alpha\}$ é convexo.*

Sejam S um aberto e $f : S \to \mathbb{R}$ diferenciável em S. Então, pode-se demonstrar que f é quasiconvexa se, e somente se, para quaisquer $x_1, x_2 \in S$ tais que $f(x_1) \leq f(x_2)$ se verifica $< \nabla f(x_2), x_1 - x_2 > \leq 0$.

Outra noção generalizada de convexidade é a seguinte:

Definição 1.14 *Sejam S um aberto e a função $f : S \to \mathbb{R}$ diferenciável em $\overline{x} \in S$. Dizemos que f é* ***pseudoconvexa em*** $\overline{x} \in S$ *se para todo $x \in S$ tal que $< \nabla f(\overline{x}), x - \overline{x} > \geq 0$ implica $f(x) \geq f(\overline{x})$. Se f for pseudoconvexa em todo $\overline{x} \in S$ dizemos que f é* ***pseudoconvexa*** *(em S).*

Estas definições podem ser relacionadas assim: Seja $f : S \to \mathbb{R}$ uma função definida num aberto, convexo, não vazio de um espaço normado. Então: f convexa $\Rightarrow$ f pseudoconvexa $\Rightarrow$ f quasiconvexa. Este conhecido resultado se encontra demonstrado em Jahn, [17], p. 94-95.

Parte I

O problema finito- dimensional

Nesta primeira parte do trabalho, estudaremos o problema (P):

$$\begin{array}{l} \text{minimizar } f_0(x) \\ \text{sujeito a } f_i(x) \leq 0, \; i = 1, ..., m \end{array} \tag{P}$$

onde os funcionais $f_0, f_1, ..., f_m$ são definidos em algum subconjunto X do $\mathbb{R}^n$.

No Capítulo 2, veremos uma condição geométrica de otimalidade local e, através de um teorema de alternativa (Teorema de Gordan Generalizado) convertemos esta condição geométrica em condições analíticas, obtendo, assim, condições necessárias dos tipos Fritz John e Karush-Kuhn-Tucker.

No Capítulo 3, estudamos um conceito generalizado de convexidade, chamado invexidade e, veremos que sob hipóteses de invexidade, as condições obtidas no Capítulo 2 resultam suficientes.

No Capítulo 4, veremos o método de deformação, o qual consiste, basicamente, em "transformarmos " um problema de difícil análise num problema de solução mais simples. Se esta "deformação " tiver boas propriedades, será possível relacionar os mínimos locais do problema original com os do problema deformado. Através do método de deformação apresentado nas seções 4.2 e 4.3; obteremos na seção 4.4 condições de otimalidade para o problema (P), sem nenhuma hipótese de convexidade (generalizada ou não) sobre os funcionais do problema (P); em 4.5 obteremos através do método de deformação um critério de otimalidade global para o problema (P).

Capítulo 2

Condições necessárias

Neste capítulo, exibimos condições necessárias de otimalidade para o problema

$$\left.\begin{array}{l} \text{minimizar } f_0(x) \\ \text{sujeito a: } f_i(x) \leq 0, \ i \in I = \{1, ..., m\} \end{array}\right\} \tag{P}$$

onde X é um subconjunto aberto e não vazio de $\mathbb{R}^n$, e as funções $f_i : X \to \mathbb{R}$ $(i = 0, 1, ..., m)$ são continuamente diferenciáveis.

Primeiramente, obtemos certas condições geométricas de otimalidade, as quais serão convertidas em condições analíticas através de um teorema de alternativa (Teorema de Gordan Generalizado).

As referências básicas para este capítulo são Contesse-Becker [14] e Mangasarian [18].

2.1 Condição geométrica de otimalidade

Definimos os seguintes conjuntos:

$$\mathcal{D}(f_i, \overline{x}) = \{v \in \mathbb{R}^n :< \nabla f_i(\overline{x}), v >< 0\}, \ i = 0, 1, ..., m.$$

Claramente estes conjuntos são cones abertos em $\mathbb{R}^n$.

Teorema 2.1 *Se $\overline{x}$ é solução local do problema (P), então*

$$\bigcap_{i\in I(\overline{x})} \mathcal{D}(f_i,\overline{x}) \bigcap \mathcal{D}(f_0,\overline{x}) = \emptyset$$

(onde $I(\overline{x})$ é o conjunto dos índices das restrições ativas em $\overline{x}$.)

Prova:

Suponha por absurdo que não e seja $y \in [\bigcap_{i\in I(\overline{x})} \mathcal{D}(f_i,\overline{x}) \bigcap \mathcal{D}(f_0,\overline{x})]$.

Como o conjunto X é aberto, então $\overline{x} + \lambda y \in X$ para $0 < \lambda < \varepsilon$ para algum $\varepsilon > 0$ suficientemente pequeno.

Seja $i \in I(\overline{x}) \cup \{0\}$.

Então: $\lim_{\lambda\downarrow 0} \dfrac{f_i(\overline{x}+\lambda y) - f_i(\overline{x})}{\lambda} =< \nabla f_i(\overline{x}), y >< 0$

Em particular, $f_i(\overline{x} + \lambda y) - f_i(\overline{x}) < 0$, $0 < \lambda < \delta_i$ para algum $\delta_i > 0$ suficientemente pequeno, $i \in I(\overline{x}) \cup \{0\}$.

Mas $f_i(\overline{x}) = 0$ para $i \in I(\overline{x})$, donde segue que $f_i(\overline{x}+\lambda y) < 0$ para $0 < \lambda < \delta_i$, $i \in I(\overline{x}) \cup \{0\}$.

Se $i \in I \backslash I(\overline{x})$, então $f_i(\overline{x}) < 0$ e pela continuidade da f segue que $f_i(\overline{x} + \lambda y) < 0$ para $0 < \lambda < \delta_i$, para algum $\delta_i > 0$ suficientemente pequeno e $i \in I \backslash I(\overline{x})$.

Seja $\delta = \min\{\varepsilon, \delta_0, \delta_1, ..., \delta_m\} > 0$.

Então, temos $f_i(\overline{x} + \lambda y) \leq 0$ e $f_0(\overline{x} + \lambda y) < f_0(\overline{x})$ para todo $0 < \lambda < \delta$, contrariando a otimalidade de $\overline{x}$. Absurdo. ■

2.2 Condições necessárias de otimalidade dos tipos Fritz John e Karush-Kuhn-Tucker

Com base no teorema anterior, damos a seguir uma condição necessária de otimalidade local, através do Teorema de Gordan Generalizado , o qual enunciamos abaixo:

Lema 2.2 *(Teorema de Gordan Generalizado)*

Sejam K um subconjunto convexo e não vazio de $\mathbb{R}^n$ e as funções $f_i : K \to \mathbb{R}$ convexas, $i = 1, ..., m$. São equivalentes as seguintes afirmações:

1. *Não existe $v \in K$ tal que $f_i(v) < 0$ $i = 1, ..., m$;*

2. *Existem* $\lambda_i \in \mathbb{R}_+$, $i = 1, ..., m$, *não todos nulos, tais que* $\sum_{i=1}^{m} \lambda_i f_i(v) \geq 0$ *para todo* $v \in K$.

(Para demonstração deste resultado, v. Contesse-Becker, [14], p. 42- 44)

Teorema 2.3 *(Fritz John)*

Se $\overline{x}$ *é solução local do problema (P), então, existem* $\overline{\lambda_0}, \overline{\lambda_i} \in \mathbb{R}$ $(i \in I)$, *não todos nulos, tais que*

$$\overline{\lambda_0}\nabla f_0(\overline{x}) + \sum_{i=1}^{m} \overline{\lambda_i}\nabla f_i(\overline{x}) = 0;$$
$$\overline{\lambda_0}, \overline{\lambda_i} \geq 0, i \in I;$$
$$\overline{\lambda_i} f_i(\overline{x}) = 0, i \in I;$$

Prova:

Se $\overline{x}$ é solução local de (P), então, pelo Teorema 2.1, não existe $v \in \mathbb{R}^n$ tal que $< \nabla f_0(\overline{x}), v >< 0$ e $< \nabla f_i(\overline{x}), v >< 0$ $(i \in I(\overline{x}))$.

Então, pelo Lema 2.2 (Teorema de Gordan Generalizado), existem $\widetilde{\lambda_0}, \widetilde{\lambda_i} \in \mathbb{R}_+$, $i \in I(\overline{x})$, não todos nulos, tais que

$$\widetilde{\lambda_0} < \nabla f_0(\overline{x}), v > + \sum_{i \in I(\overline{x})} \widetilde{\lambda_i} < \nabla f_i(\overline{x}), v > \geq 0$$

para todo $v \in \mathbb{R}^n$ ou seja,

$$\widetilde{\lambda_0}\nabla f_0(\overline{x}) + \sum_{i \in I(\overline{x})} \widetilde{\lambda_i}\nabla f_i(\overline{x}) = 0.$$

Basta, então, tomarmos

$$\overline{\lambda_0} = \widetilde{\lambda_0}, \overline{\lambda_i} = \begin{cases} \widetilde{\lambda_i}, & \text{se } i \in I(\overline{x}) \\ 0, & \text{se } i \in I \backslash I(\overline{x}) \end{cases}$$

■

Observe que no Teorema 2.3 nada garante que o multiplicador associado à função objetivo é diferente de zero. Assim, se este multiplicador for nulo, toda a informação contida na derivada do funcional objetivo será perdida e as condições de Fritz John perderão seu valor na busca de solução ótima.

Veremos a seguir que sob certas condições de regularidade sobre as restrições do problema (P), esta situação pode ser contornada.

Assuma a seguinte condição de regularidade:

$$\bigcap_{i\in I(\overline{x})} \mathcal{D}(f_i,\overline{x}) \neq \emptyset \qquad (**)$$

Teorema 2.4 *(Karush-Kuhn-Tucker)*

Seja $\overline{x}$ uma solução local de (P) e suponha que a hipótese de regularidade (**) seja satisfeita. Então, existem $\widehat{\lambda_i} \in \mathbb{R}$, $i \in I$, tais que

$$\nabla f_0(\overline{x}) + \textstyle\sum_{i=1}^{m} \widehat{\lambda_i} \nabla f_i(\overline{x}) = 0;$$
$$\widehat{\lambda_i} \geq 0, i \in I;$$
$$\widehat{\lambda_i} f_i(\overline{x}) = 0, \; i \in I.$$

Prova:

Note que pelo Teorema 2.3, existem $\overline{\lambda_0}, \overline{\lambda_i} \in \mathbb{R}_+$ $(i \in I(\overline{x}))$ tais que

$$\overline{\lambda_0} \nabla f_0(\overline{x}) + \textstyle\sum_{i\in I(\overline{x})} \overline{\lambda_i} \nabla f_i(\overline{x}) = 0.$$

Suponha por absurdo que se tenha $\overline{\lambda_0} = 0$.

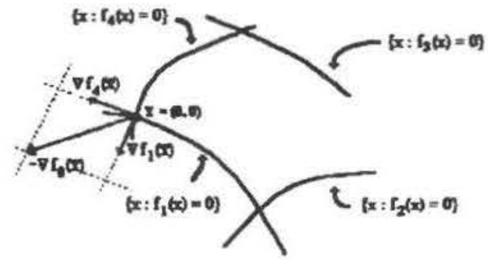
Então, podemos afirmar que

$$\sum_{i\in I(\overline{x})} \overline{\lambda_i} < \nabla f_i(\overline{x}), v >= 0$$

para todo $v \in \mathbb{R}^n$.

Logo, pelo, Lema 2.2, não existe $v \in \mathbb{R}^n$ tal que $< \nabla f_i(\overline{x}), v >< 0$ para todo $i \in I(\overline{x})$.

Mas a condição de regularidade (**) afirma que existe um $v_0 \in \mathbb{R}^n$ tal que $< \nabla f_i(\overline{x}), v >< 0$ para todo $i \in I(\overline{x})$, contrariando a afirmação anterior.



Interpretação geométrica das condições K-K-T

Figure 2-1:

O resultado desejado é obtido fazendo $\widehat{\lambda_i} = \frac{\overline{\lambda_i}}{\overline{\lambda_0}}$. ■

A condição de Karush- Kuhn- Tucker (ou abreviadamente, KKT) admite a seguinte interpretação geométrica, a qual ilustramos na figura 2-1:

Admitamos, no problema (P), $m = 4$ e $n = 2$.

Na figura, consideramos os conjuntos de nível $E_0(f_i) = \{x : f_i(x) = 0\}$associados a cada uma das restrições. Neste caso, o conjunto das restrições ativas em $\overline{x} = (0,0)$ é dado por $I(\overline{x}) = \{1,4\}$.

Os gradientes são, naturalmente, ortogonais aos seus conjuntos de nível no ponto $\overline{x}$ e apontam para a direção de máximo crescimento de f_i a partir de $\overline{x}$.

Se $\overline{x}$ é uma solução factível e regular e é mínimo local para o problema, então o teorema de Karush- Kuhn- Tucker garante a existência de $\lambda_1, \lambda_4 \geq 0$ tais que

$$-\nabla f_0(\overline{x}) = \lambda_1 \nabla f_1(\overline{x}) + \lambda_4 \nabla f_4(\overline{x})$$

ou seja, $\nabla f_0(\overline{x})$ aponta para uma direção oposta à determinada pelos vetores $\nabla f_1(\overline{x}), \nabla f_4(\overline{x})$.

Fazemos ainda as seguintes observações:

1. Um ponto $\overline{x} \in \mathbb{F}$ que satisfaz as condições de Fritz John para algum $\overline{\lambda_0}, \overline{\lambda_i} \in \mathbb{R}$, $i \in I$, é chamado um **extremo** de (P).

2. Existem outros critérios de regularidade que garantem que o multiplicador do funcional objetivo, $\overline{\lambda_0}$, é não nulo. O critério de regularidade mais conhecido é o de **Slater**: se existe $\widehat{x}$ factível para (P) tal que $f_i(\widehat{x}) < 0$ para todo $i \in I(\overline{x})$, e $\overline{x}$ é solução local de (P), então valem as condições Fritz John , com $\overline{\lambda_0} > 0$. Outro critério de regularidade bastante conhecido é o de **Mangasarian-Fromovitz**: se os vetores $\nabla f_i(\overline{x})$, $i \in I(\overline{x})$ são linearmente independentes e $\overline{x}$ é solução local de (P), então valem as condições Fritz John, com $\overline{\lambda_0} > 0$ (v. Contesse-Becker, [14]).

3. Para o problema (P), definimos a **função Lagrangeana** $\mathcal{L} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$ onde $\mathcal{L}(x, \overline{\lambda_0}, \overline{\lambda}) = \overline{\lambda_0} f_0(x) + < \overline{\lambda}, f(x) >$, onde $f = (f_1, ..., f_m)$. Deste modo, a primeira equação do Teorema 2.3 pode ser rescrita como $\nabla_x \mathcal{L}(\overline{x}, \overline{\lambda_0}, \overline{\lambda}) = 0$.

Observação 2.5 *Em [18], os autores demonstram os teoremas de Fritz John e Karush Kuhn Tucker diretamente através do Teorema de Separação de Hahn- Banach. Preferimos demonstrar esses teoremas através do Teorema de Gordan Generalizado pois utilizaremos esta mesma idéia na obtenção de condições necessárias de otimalidade para o Problema de Tempo Contínuo (v. Capítulo 5).*

Capítulo 3

Condições suficientes para o programa invexo

Veremos a seguir que para uma outra classe de funcionais, um pouco mais ampla que a das funções convexas- a saber- as funções invexas, introduzidas por Hanson em 1981 em [16], as condições de Karush- Kuhn- Tucker ainda resultam suficientes. Para isto, definimos:

Definição 3.1 *Sejam Ω um subconjunto aberto, não vazio de $\mathbb{R}^n$ e a função $f : \Omega \to \mathbb{R}$ diferenciável. A função f é dita* ***invexa*** *em $u \in \Omega$ se existe uma função $\eta : \Omega \times \Omega \to \mathbb{R}^n$ tal que*

$$f(x) - f(u) \geq < \nabla f(u), \eta(x,u) >$$

para cada $x \in \Omega$. Dizemos, simplesmente, que f é invexa , se for invexa em todo $u \in \Omega$.

Note que se f é uma função convexa diferenciável, em particular, f é invexa: basta tomarmos $\eta(x,u) = x - u$. Além disso, as noções de convexidades generalizadas que foram mencionadas até agora podem ser relacionadas como segue: se f é uma função diferenciável, temos que convexidade implica pseudoconvexidade e pseudoconvexidade implica quasiconvexidade. Por outro lado, pseudoconvexidade implica invexidade (como veremos a seguir).

O exemplo seguinte mostra que nem toda função invexa é convexa (o que mostra que a classe das funções invexas é estritamente maior que a das funções convexas).

Exemplo 3.2 *Em [16] é apresentado o seguinte exemplo de programa não convexo cujos funcionais satisfazem a Definição 3.1 para uma mesma função η :*

$$\begin{aligned}
&\text{minimizar } x_1 - \sin x_2 \\
&\text{sujeito a:} \\
&\ \sin x_1 - 4\sin x_2 \leq 0, \\
&\ 2\sin x_1 + 7\sin x_2 + x_1 - 6 \leq 0, \\
&\ 2x_1 + 2x_2 - x \leq 0, \\
&\ 4x_1^2 + 4x_2^2 - 9 \leq 0, \\
&\ -\sin x_1 \leq 0, \\
&\ -\sin x_2 \leq 0, \\
&\ (x_1, x_2) \in \mathbf{R}^2.
\end{aligned}$$

É clara a natureza não convexa dos funcionais deste problema. No entanto, todos eles satsfazem a Definição 3.1, com $\eta(y,x)$ dada por

$$\eta(y,x) = (\frac{\sin y_1 - \sin x_1}{\cos x_1}, \frac{\sin y_2 - \sin x_2}{\cos x_2})$$

onde $x = (x_1, x_2)$, $y = (y_1, y_2)$.

Deste modo, vemos que o conjunto das funções invexas é estritamente maior que o das funções convexas.

Antes de enunciarmos e demonstrarmos o teorema das condições suficientes de otimalidade para o programa invexo (isto é, quando se tem todos os funcionais em (P) invexos, para uma mesma função η), necessitaremos dos seguintes dois lemas:

Lema 3.3 *Uma função f é invexa em Ω se e somente se todo ponto estacionário de f é ponto de mínimo global.*

Prova:

Suponha f invexa em Ω e seja u um ponto estacionário de f ou seja,$\nabla f(u) = 0$. Da definição de invexidade, segue que

$$f(x) - f(u) \geq < 0, \eta(x,u) >= 0$$

e, portanto, $f(u) \leq f(x)$ para todo $x \in \Omega$ e, assim, u é mínimo global de f.

Por outro lado, suponha que todo ponto estacionário de f é ponto de mínimo global.

Agora, defina:

$$\eta(x,u) = \begin{cases} \dfrac{(f(x) - f(u))}{||\nabla f(u)||^2}.\nabla f(u) \text{ se } \nabla f(u) \neq 0 \\ 0, \text{ caso contrário} \end{cases}$$

Se $\nabla f(u) = 0$, então $< \nabla f(u), \eta(x,u) >= 0 \leq f(x) - f(u)$, pois u é ponto de mínimo global.

Se $\nabla f(u) \neq 0$, então $< \nabla f(u), \eta(x,u) >= f(x) - f(u)$

Assim, em qualquer caso, a função f é invexa. ■

Notemos que do Lema 3.3 segue que toda função pseudoconvexa é invexa. De fato, seja f uma função pseudoconvexa e seja $\overline{x}$ tal que $\nabla f(\overline{x}) = 0$. Neste caso, trivialmente, temos $< \nabla f(\overline{x}), x - \overline{x} > \geq 0$ para todo x e, portanto, da definição de pseudoconvexidade, temos que $f(x) \geq f(\overline{x})$, para todo x.

Lema 3.4 *Se para cada $i = 1, ..., m$ temos $\lambda_i \geq 0$ e as funções f_i são invexas para uma mesma função η, então a função $\sum_{i=1}^{m} \lambda_i f_i$ é invexa.*

Prova:

Da definição de invexidade segue que

$$f_i(x) - f_i(u) \geq < \nabla f_i(u), \eta(x,u) >, i = 1, ..., m$$

Como $\lambda_i \geq 0$ para todo $\lambda_i \geq 0$, temos que

$$\lambda_i(f_i(x) - f_i(u)) \geq < \lambda_i \nabla f_i(u), \eta(x,u) >, i = 1, ..., m$$

Somando sobre os índices i obtemos

$$\textstyle\sum_{i=1}^{m} \lambda_i f_i(x) - \lambda_i f_i(u) \geq \sum_{i=1}^{m} < \lambda_i \nabla f_i(u), \eta(x,u) >$$

E, portanto,

$$\sum_{i=1}^{m} \lambda_i f_i(x) - \sum_{i=1}^{m} \lambda_i f_i(u) \geq < \nabla[\sum_{i=1}^{m} \lambda_i f_i(u)], \eta(x,u) >$$

Logo, a função $\sum_{i=1}^{m} \lambda_i f_i$ é invexa. ■

Agora, estamos em condição de enunciar e demonstrar o seguinte

Teorema 3.5 *(condições suficientes de otimalidade para o programa invexo) Suponha que x^* seja um ponto factível para (P) e que as funções f_0, f_i são invexas para uma mesma função η e que existam multiplicadores $\lambda_i \in \mathbb{R}$, $i = 1, ..., m$ tal que as condições de Karush- Kuhn- Tucker são satisfeitas em x^* . Então x^* é um ponto de mínimo global para (P).*

Prova:

Pelo Lema 3.4, a função $f_0 + \sum_{i=1}^{m} \lambda_i f_i$ é invexa em x^* e das condições de Karush- Kuhn-Tucker, segue que x^* é ponto estacionário de $f_0 + \sum_{i=1}^{m} \lambda_i f_i$. Então, pelo Lema 3.3 , x^* é ponto de mínimo global de $f_0 + \sum_{i=1}^{m} \lambda_i f_i$. Deste modo,

$$f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \geq f_0(x^*) + \sum_{i=1}^{m} \lambda_i f_i(x^*)$$

para todo ponto factível para (P).

Mas $\lambda_i f_i(x*) = 0$ para todo índice i e, portanto,

$$f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \geq f_0(x^*)$$

donde resulta

$$f_0(x) \geq f_0(x^*)$$

pois $\lambda_i f_i(x) \leq 0$ para todo x factível para (P).

Logo, x^* é solução global de (P). ■

Capítulo 4

O método de deformação e condições suficientes de otimalidade

4.1 Introdução

Neste capítulo, consideraremos uma família parametrizada de problemas do seguinte tipo:

$$\begin{aligned}
&\text{minimizar } f(x;\lambda)\\
&\text{sujeito a}\\
&f_i(x;\lambda) \leq 0, i = 1, ..., m,\\
&x \in \mathbb{R}^N
\end{aligned} \tag{P(λ)}$$

onde o parâmetro $\lambda \in [0, 1]$.

Deste modo, é possível "transformar " um problema $(P(0))$ de difícil análise, num problema $(P(1))$,de mais fácil investigação.

Grosso modo, se para cada $\lambda \in [0, 1]$ o problema $(P(\lambda))$ admite um extremo uniformememente isolado (com respeito a λ) e este extremo minimiza o problema inicial, então também minimiza o problema deformado- este é o resultado central deste capítulo e será enunciado mais cuidadosamente na Seção 4.3.

As referências básicas para esta seção são [6], [7], [9] e [11].

4.2 Deformação de funcionais Lipschitz

Nesta seção, estudamos um resultado sobre a deformação de funcionais Lipschitz, o qual será usado no estudo da família de problemas (P(λ)). Antes, porém, definimos:

Definição 4.1 *Sejam $B(R)$ a bola fechada de raio R em $\mathbb{R}^N$ e $f(x;\lambda)$ um funcional definido em $B(R) \times [0,1]$. Dizemos que $f(x;\lambda)$ é uma* ***deformação não degenerada*** *se satisfaz as seguintes três condições:*

(a) *$f(x;\lambda)$ é contínua em $B(R) \times [0,1]$ e localmente Lipschitz com respeito a x, para todo $\lambda \in [0,1]$.*

(b) *A multifunção $\partial_x f(x;\lambda) : B(R) \times [0,1] \rightrightarrows \mathbb{R}^N$ é semicontínua superior.*

(c) *Para todo $\lambda \in [0,1]$ o funcional $f(x;\lambda)$ tem um único ponto crítico, $x(\lambda)$, continuamente dependente de $\lambda \in [0,1]$.*

Observação 4.2 *Lembramos que um ponto x^* é dito um* ***ponto crítico*** *do funcional f se $0 \in \partial f(x^*)$.*

Este tipo de deformação preserva a propriedade de um extremo ser mínimo local. É o que nos diz o próximo resultado, cuja demonstração omitiremos [1] e pode ser encontrada em [8].

Lema 4.3 *(deformação de funcionais) Seja $f(x;\lambda)$ uma deformação não degenerada. Se $x_0 = x(0)$ é mínimo local de $f(x;0)$ então $x_1 = x(1)$ é mínimo local de $f(x;1)$.*

4.3 Deformação de problemas de programação não linear

Nesta seção, enunciaremos um teorema que nos permite obter condições suficientes de otimalidade para o problema clássico de programação não linear, o qual formulamos abaixo:

$$\begin{aligned} &\text{minimizar } f(x) \\ &\text{sujeito a: } f_i(x) \leq 0, i = 1, ..., m, x \in \mathbb{R}^N \end{aligned} \tag{P}$$

[1] esta demonstração é demasiado longa e técnica o que justifica tal omissão

onde os funcionais f, f_i são supostos continuamente diferenciáveis.

Antes de enunciá-lo, porém, definimos

Definição 4.4 *A família parametrizada de problemas*

$$\begin{array}{l} \text{minimizar } g(x;\lambda) \\ \text{sujeito a: } g_i(x;\lambda) \leq 0, i = 1,...,m; x \in \mathbb{R}^N \end{array} \qquad \text{(P(}\lambda\text{))}$$

*onde $\lambda \in [0,1]$ é chamada uma **deformação não singular** se satisfaz as seguintes três condições:*

1. *As funções $g(x;\lambda)$, $g_i(x;\lambda), \nabla_x g(x;\lambda), \nabla_x g_i(x;\lambda)$ são contínuas em $\mathbb{R}^N \times [0,1]$, para todo $i = 1,...,m$.*

2. *Para cada $\lambda \in [0,1]$, o problema $(P(\lambda))$ tem um extremo $x(\lambda)$, continuamente dependente de $\lambda \in [0,1]$ e, uniformemente isolado com respeito a λ (ou seja, existe um $\rho > 0$ tal que para todo $\lambda \in [0,1]$, a bola centrada em $x(\lambda)$ não contém nenhum extremo de (P_λ), além de $x(\lambda)$).*

3. *Para cada $\lambda \in [0,1]$ os vetores $\nabla_x g_i(x(\lambda);\lambda), i = 1,...,m$ são linearmente independentes e, para todo $\lambda \in [0,1]$ existe um vetor $s(\lambda) \in \mathbb{R}^N$, tal que $< \nabla_x g_i(x(\lambda);\lambda), s(\lambda) >< 0$ para todo $i \in I(x(\lambda))$ (onde $I(x(\lambda))$ é o conjunto de índices de restrições ativas em $x(\lambda)$).*

Teorema 4.5 *Seja $(P(\lambda))$ como definida acima, uma deformação não singular. Se $x_0 = x(0)$ é ponto de mínimo local do problema $(P(0))$, então, $x_1 = x(1)$ é mínimo local do problema $(P(1))$.*

A demonstração do Teorema 4.5, que é demasiado longa e técnica- será omitida neste trabalho e pode ser encontrada em [6], p. 919-921. Ao invés de demonstrarmos este resultado, enunciaremos e demonstraremos o caso mais simples em que se admite que o problema $(P(\lambda))$ possui um único extremo para todo $\lambda \in [0,1]$, o que será feito com o auxílio do Lema 4.3.

Para isto, consideremos o problema

$$\begin{array}{l} \text{minimizar } f(x) \\ \text{sujeito a: } g_i(x) \leq 0, i = 1,...,m, x \in \mathbb{R}^N \end{array} \qquad \text{(P')}$$

onde as funções $f, g_i : \mathbb{R}^N \to \mathbb{R}$ são localmente Lipschitz, ou seja, não supomos nenhuma hipótese de diferenciabilidade no problema.

Neste caso, dizemos que um ponto factível para (P'), x^*, é um **extremo** se existem $\mu, y_1, ..., y_m \geq 0$ não todos nulos tais que

$$0 \in \mu\partial f(x^*) + \textstyle\sum_{i=1}^m y_i \partial g_i(x^*)$$
$$y_i g_i(x^*) = 0, i = 1, ..., m$$

Pode-se, ainda, definir extremo de (P') em termos do cone normal ao conjunto $Q = \{x \in \mathbb{R}^N : g_i(x) \leq 0, i = 1, ..., m, x \in \mathbb{R}^N\}$, $N_Q(x^*)$: neste caso, dizemos que x^* é um extremo de (P') se

$$0 \in \partial f(x^*) + N_Q(x^*).$$

Agora, definimos deformação não degenerada (em contexto não diferenciável) como segue

Definição 4.6 *A família parametrizada (em λ) de problemas*

$$\begin{aligned} &\text{minimizar } f(x;\lambda) \\ &\text{sujeito a: } g(x;\lambda) \leq 0, x \in \mathbb{R}^N \end{aligned} \qquad (\mathrm{P}(\lambda))$$

*(onde $\lambda \in [0,1]$) é chamada uma **deformação não degenerada** se satisfaz as seguintes quatro condições:*

1. *Os funcionais f, g são contínuos em $\mathbb{R}^N \times [0,1]$ e são localmente lipschitzianos com respeito a x, para todo $\lambda \in [0,1]$.*

2. *As multifunções*

$$\partial_x f(x;\lambda) : \mathbb{R}^N \times [0,1] \rightrightarrows \mathbb{R}^N$$
$$\partial_x g(x;\lambda) : \mathbb{R}^N \times [0,1] \rightrightarrows \mathbb{R}^N$$

são semicontínuas superiores, para cada $\lambda \in [0,1]$.

3. *Para todo x tal que $g(x;\lambda) = 0$, a derivada generalizada $g^0(x,v;\lambda)$ coincide com a derivada direcional clássica na direção v, $g'(x,v;\lambda)$ e $0 \notin \partial_x g(x;\lambda)$.*

4. *Para todo $\lambda \in [0,1]$,o problema $(P(\lambda))$ tem um único extremo $x(\lambda)$, continuamente dependente de λ.*

A seguir, veremos que uma tal deformação preserva a propriedade de um extremo ser mínimo local: é o que afirma o seguinte

Teorema 4.7 *Seja $(P(\lambda))$ como definida anteriomente, uma deformação não degenerada. Se $x_0 = x(0)$ é ponto de mínimo local do problema $(P(0))$, então $x_1 = x(1)$ é ponto de mínimo local do problema $(P(1))$.*

Prova:

Por simplicidade, assumiremos $f(0;\lambda) \equiv g(0,\lambda) \equiv 0$ e $x(\lambda) \equiv 0$, para todo $\lambda \in [0,1]$.

Sejam

$$d_{Q(\lambda)}(x) = \inf_{y \in Q(\lambda)} |x - y| \tag{1}$$

onde $Q(\lambda) = \{x \in R^N : g(x;\lambda) \leq 0\}$ e, para $k \in \mathbb{R}_+$ definimos

$$F(x,k,\lambda) = f(x;\lambda) + kd_{Q(\lambda)}(x) \tag{2}$$

onde $x \in \mathbb{R}^N, \lambda \in [0,1]$.

Veremos que, para k suficientemente grande, existe uma bola $B(\rho)$ de raio $\rho > 0$ tal que $F(x,k,\lambda)$ satisfaz as hipóteses do Lema 4.3 em $B(\rho)$.

Para isto, é suficiente mostrar que para k suficientemente grande a origem é mínimo local de $F(x,k,0)$ e é um ponto crítico de $F(x,k,\lambda)$, uniformemente isolado com respeito a λ.

Seja k_1 a constante de Lipschitz de $f_0 = f(\cdot,0)$ na bola unitária.

Pela Proposição 2.4.3, p. 51, [13], se $\rho < 1$ e $k > k_1$, os pontos de mínimo de $F(x,k,0)$ em $B(\rho)$ estão no conjunto $Q_0 \cap B(\rho)$,onde $Q_0 = Q(0)$.

Mas, para ρ suficientemente pequeno, a origem é o único ponto de mínimo de $F(x,k,0)$ em $Q_0 \cap B(\rho)$.

De fato, caso contrário, para cada $n \in \mathbb{N}$, existe $x_n \neq 0$ ponto crítico de $F(x,k,0)$, ou seja, $0 \in \partial f(x_n,0) + k\partial d_{Q(0)}(x_n)$.

Mas $\partial d_{Q(0)}(x_n) \subset N_{Q(0)}(x_n)$.

Pois do contrário, existiria um $\xi_0 \in \partial d_{Q(0)}(x_n)$ tal que $\xi_0 \notin N_{Q(0)}(x_n)$, ou seja, existe um $v_0 \in T_{Q(0)}(x_n)$ tal que

$$0 << \xi_0, v_0 >\leq d^0_{Q(0)}(x_n, v_0)$$

Assim, $d^0_{Q(0)}(x_n, v_0) > 0$ contrariando $v_0 \in T_{Q(0)}(x_n)$. Absurdo.

Deste modo, temos

$$0 \in \partial f(x_n, 0) + k\partial d_{Q(0)}(x_n) \subset \partial f(x_n, 0) + N_{Q(0)}(x_n)$$

e $x_n \neq 0$ é extremo de (P(0)), o que é absurdo.

Logo, para ρ suficientemente pequeno, a origem é o único ponto de mínimo de $F(x, k, 0)$ em $Q_0 \cap B(\rho)$.

Portanto, a origem é mínimo local de $F(x, k, 0)$ para $k > k_1$.

Resta provar que para k suficientemente grande, $F(x, k, \lambda)$ não admite ponto crítico não nulos, com normas arbitrariamente pequenas.

De fato, suponha que $F(x, k, \lambda)$ admita um ponto crítico $x_0 \in \partial Q(\lambda)\backslash\{0\}$, com $|x_0| << 1$. Neste caso, temos:

$$\begin{aligned} 0 &\in \partial_x(f(x_0, \lambda) + kd_{Q(\lambda)}(x_0)) \subset \partial_x f(x_0, \lambda) + k\partial_x d_{Q(\lambda)}(x_0) \\ &\subset \partial_x f(x_0, \lambda) + \cup_{\mu \geq 0}\mu\partial d_{Q(\lambda)}(x_0, \lambda) = \partial_x f(x_0, \lambda) + N_{Q(\lambda)}(x_0) \end{aligned}$$

e, portanto,

$$0 \in \partial_x f(x_0, \lambda) + N_{Q(\lambda)}(x_0).$$

(a igualdade anterior segue da Proposição 2.4.2, p. 51, de [13]).

Deste modo o problema $(P(\lambda))$ admite um extremo $x_0 \neq 0$. Absurdo.

Agora, suporemos que exista um $x_0 \in \operatorname{int} Q(\lambda)\backslash\{0\}$, com norma arbitrariamente pequena,

$x_0 \in \partial_x F(x_0, k, \lambda)$. Neste caso, $x_0 \in \partial_x f(x_0, \lambda)$, o que contraria as hipóteses do teorema.

Portanto, o funcional $F(x, k, \lambda)$ não possui pontos críticos de norma arbitrariamente pequenas em $Q(\lambda)$.

Veremos que, para k suficientemente grande, $F(x, k, \lambda)$ não possui pontos críticos de norma arbitrariamente pequenas em $\mathbb{R}^N \backslash Q(\lambda)$.

Seja $M = \text{co}\{(\bigcup_{\mu \geq 0} \mu \partial_x g(0, \lambda)) \cap \partial B\}$.

Pode-se provar (e isto é feito em [9], p. 2693-2694) que existem $\varepsilon_0, \varepsilon_1 > 0$ tal que $\partial_x d_{Q(\lambda)}(x) \subset V(\varepsilon_0, M)$, para $x \in \mathbb{R}^N \backslash Q(\lambda), |x| < \varepsilon_1$. (onde $V(\varepsilon_0, M) = \{x \in \mathbb{R}^N : |y - x| \leq \varepsilon_0, y \in M\}$).

Sejam agora,

$$
\begin{gathered}
a = \min_{y \in V(\varepsilon_0, M)} |y| \\
A = \max_{\substack{y \in \partial_x f(x,\lambda) \\ x \in B}} |y| \\
k_2 = \frac{A}{a}
\end{gathered}
$$

Então, para todo $x \notin Q(\lambda), |x| < \varepsilon_1, k > k_2$, temos

$$
\begin{aligned}
\min_{y \in \partial_x F(x,k,\lambda)} |y| &= \min_{y \in \partial_x (f(x,\lambda) + k d_{Q(\lambda)}(x))} |y| \geq \min_{y \in (\partial_x f(x,\lambda) + k \partial_x d_{Q(\lambda)}(x))} |y| \geq \\
&\geq \min_{y \in (B(A) + kV(\varepsilon_0, M))} |y| \geq ka - A > 0
\end{aligned}
$$

e, portanto, temos que $0 \notin \partial_x F(x, k, \lambda)$.

Assim, para $k \geq k_0 = \max\{k_1, k_2\}$a origem é ponto crítico de $F(x, k, \lambda)$ isolado com respeito a $\lambda \in [0, 1]$.

Deste modo, para todo $k \geq k_0$, $F(x, k, \lambda)$ é uma deformação não degenerada (no sentido da definição 4.1). Por hipótese, temos que a origem é mínimo local de $F(x, k, 0)$ e, portanto, pelo Lema 4.3, é mínimo local de $F(x, k, 1)$.

Seja $k > k_0$. Então, existe um $R > 0$ tal que

$$
F(0, k, 1) = f(0, 1) + k d_{Q(1)}(0) = 0 \leq F(x, k, 1) = f(x, 1) + k d_{Q(1)}(x)
$$

para todo $x \in \mathbb{R}^N$, com $|x| < R$.

Em particular, se $x \in Q(1)$ tal que $|x| < R$, $d_{Q(1)}(x) = 0$ e portanto,

$$0 = f(0,1) \leq f(x,1)$$

para todo x factível para (P(1)), $|x| < R$, ou seja, a origem é solução local de (P(1)). ∎

Agora, estudaremos uma deformação de problemas diferenciáveis, com várias restrições de desigualdade:

Definição 4.8 *A família parametrizada de problemas*

$$\begin{aligned} &\textit{minimizar } f(x;\lambda) \\ &\textit{sujeito a: } g_i(x;\lambda) \leq 0, i = 1, ..., m; x \in \mathbb{R}^N \end{aligned} \qquad (P(\lambda))$$

onde $\lambda \in [0,1]$ *é chamada uma* ***deformação não degenerada*** *se satisfaz:*

1. *As funções* $f(x;\lambda), g_i(x;\lambda), \nabla_x f(x;\lambda), \nabla_x g_i(x;\lambda)$ *são contínuas em* $\mathbb{R}^N \times [0,1]$, *para todo* $i = 1, ..., m$.

2. *Seja* $Q(\lambda) = \{x \in \mathbb{R}^N : g_i(x;\lambda) \leq 0, i = 1, ..., m\}$. *Se* $x \in \partial Q(\lambda)$ *então os vetores* $\nabla_x g_i(x;\lambda)$ *são positivamente independentes, para* $i \in I_\lambda(x)$ *(onde* $I_\lambda(x)$ *é o conjunto dos índices ativos em* x, *para o problema* $(P(\lambda))$.

3. *Para todo* $\lambda \in [0,1]$ *o problema* $(P(\lambda))$ *tem um único extremo,* $x(\lambda)$, *continuamente dependente de* λ.

O Teorema 4.9, o qual enunciamos e demonstramos a seguir afirma que uma tal deformação preserva a propriedade de um extremo ser mínimo local. De fato:

Teorema 4.9 *Suponha que a família parametrizada de problemas,* $(P(\lambda))$ *como definida acima seja uma deformação não degenerada. Seja* $x_0 = x(0)$ *ponto de mínimo local do problema* $(P(0))$. *Então o ponto* $x_1 = x(1)$ *é mínimo local do problema* $(P(1))$.

Prova:

Assuma por simplicidade que $f(0,\lambda) \equiv 0 \equiv g_i(0,\lambda)$ $(\lambda \in [0,1], i = 1, ..., m)$ e que $x(\lambda) \equiv 0$ $(\lambda \in [0,1])$.

Note que para cada $\lambda \in [0,1]$,o problema $(P(\lambda))$ é equivalente ao seguinte problema:

$$\text{minimizar } f(x;\lambda) \tag{1}$$

$$\max_{i=1,\dots,m} g_i(x;\lambda) = g(x;\lambda) \leq 0 \tag{2}$$

Verificaremos que a família de problemas acima satisfaz as hipóteses do Teorema 4.7 . Para isto, basta provar que a origem é o único extremo de (1)-(2) e que a multifunção

$$\partial_x g(x;\lambda) : Q(\lambda) \times [0,1] \rightrightarrows \mathbb{R}^N$$

é semicontínua superior e que, além disso, é tal que para todo $\lambda \in [0,1]$ a restrição (2) satisfaz as condições de regularidade do Teorema 4.7.

Como para todo $\lambda \in [0,1]$ a origem é um extremo de $(P(\lambda))$, então existem multiplicadores $\mu \in \mathbb{R}_+, y = (y_1, \dots, y_m) \in \mathbb{R}_+^m$, não simultaneamente nulos, tais que

$$\mu \nabla_x f(0;\lambda) + \sum_{i=1}^{m} y_i \nabla_x g_i(0;\lambda) = 0 \tag{3}$$

Como os vetores $\nabla_x g_i(0;\lambda)$ são positivamente independentes, segue que $\mu \neq 0$. Neste caso, podemos rescrever (3) como

$$\nabla_x f(0;\lambda) + \sum_{i=1}^{m} \tau_i \nabla_x g_i(0;\lambda) = 0 \tag{4}$$

onde $\tau_i = \dfrac{y_i}{\mu}$.

Pode-se provar (v. Clarke, [13], p. 57) que o cone normal a $Q(\lambda)$ na origem é dado por

$$N_{Q(\lambda)}(0) = \{\sum_{i=1}^{m} \mu_i \nabla_x g_i(0;\lambda) : \mu_i \geq 0, i = 1, \dots, m\} \tag{5}$$

E portanto, de (4) e (5) segue que

$$0 \in \nabla_x f(0;\lambda) + N_{Q(\lambda)}(0)$$

e, logo, a origem é um extremo do problema (1)-(2).

Agora, suponha que para um certo $\lambda \in [0,1]$, o problema (1)-(2) admita um extremo $x^* \in Q(\lambda)\backslash\{0\}$. Assim,

$$0 \in \nabla_x f(x^*;\lambda) + N_{Q(\lambda)}(x^*) \tag{6}$$

Novamente, o cone normal a $Q(\lambda)$ em x^* é dado por

$$N_{Q(\lambda)}(x^*) = \{\sum_{i=1}^{m} \mu_i \nabla_x g_i(x^*;\lambda) : \mu_i \geq 0, i = 1,...,m\} \tag{7}$$

e de (6) e (7) segue que existem $\mu_i^* \geq 0, i \in I_\lambda(x^*)$ tais que

$$\nabla_x f(x^*;\lambda) + \sum_{i \in I_\lambda(x^*)} \mu_i^* \nabla_x g_i(x*;\lambda) = 0$$

e, portanto, $x^* \neq 0$ é extremo do problema $(P(\lambda))$. Absurdo.

Logo, para todo $\lambda \in [0,1]$ a origem é o único extremo de (1)-(2).

Além disso, pode-se provar, (v. Clarke, [13], p. 47) que

$$\partial_x g(x;\lambda) = \mathrm{co}\{\nabla g_i(x;\lambda) : i \in I_\lambda(x)\} \tag{8}$$

para todo $x \in Q(\lambda)$ e como a função $I_\lambda(x)$ é semicontínua superior em $Q(\lambda) \times [0,1]$, segue que a multifunção $\partial_x g(x;\lambda)$ é semicontínua superior em $Q(\lambda) \times [0,1]$.

Notemos, ainda, que as $g_i(x;\lambda)$ são continuamente diferenciáveis, e, portanto, regulares.

Temos que, para cada $x \in \partial Q(\lambda)$, $0 \notin \partial_x g(x;\lambda)$, pois, do contrário, teríamos

$$0 = \sum_{i \in I_\lambda(x)} \beta_i \nabla g_i(x;\lambda)$$

para algum $\beta_i \geq 0$, $\sum_{i \in I_\lambda(x)} \beta_i = 1$. Mas os vetores$\nabla g_i(x;\lambda)$ $(i \in I_\lambda(x))$ são positivamente independentes, donde segue que $\beta_i = 0$, para cada $i \in I_\lambda(x)$, o que é absurdo (pois contraria $\sum_{i \in I_\lambda(x)} \beta_i = 1$).

Logo, todas as hipóteses do Teorema 4.7 são satisfeitas e, portanto, a origem é solução do problema $(P(1))$. ■

4.4 Condições suficientes via Método de Deformação

Neste capítulo, consideraremos o seguinte problema de programação não linear:

$$\begin{aligned} &\text{minimizar } g_0(x) \\ &\text{sujeito a: } g_i(x) \leq 0 (x \in \mathbb{R}^N; i = 1, ..., m) \end{aligned} \tag{P}$$

onde as funções g_0, g_i são continuamente diferenciáveis.

É fato que os métodos clássicos de investigação de condições suficientes exigem que os funcionais do problema (P) sejam duas vezes continuamente diferenciáveis (métodos de segunda ordem) ou alguma hipótese de convexidade sobre os funcionais do problema.

Através do método de deformação, Bobylev [7] obtém novas condições suficientes de otima lidade para o problema (P), sem nenhuma hipótese de convexidade (generalizada ou não).

Teorema 4.10 *(condições suficientes de otimalidade) Assuma no problema (P) que* $m = N$. *Sejam* $a_i = \nabla g_i(0)$ $(i = 0, 1, ..., m)$,*e suponha que a origem seja um extremo de (P), cujos multiplicadores correspondentes* $\mu_0, y_i^0 (i = 1, ..., m)$ *são positivos e que os vetores* a_i $(i = 1, ..., m)$ *são linearmente independentes. Então a origem é mínimo local de (P).*

Prova:

Consideremos a seguinte família parametrizada de problemas:

$$\min(1 - \lambda)g_0(x) + \lambda < a_0, x > \tag{1}$$

$$(1 - \lambda)g_i(x) + \lambda < a_i, x > \leq 0 (i = 1, ..., m; x \in \mathbb{R}^N) \tag{2}$$

onde $\lambda \in [0, 1]$.

Sem perda de generalidade, podemos assumir

$$g_0(0) = 0 = g_i(0) \ , \ i = 1, ..., m$$

Temos que a origem é um extremo de (1),(2), para cada $\lambda \in [0, 1]$.

De fato, sejam $\alpha_0, z_i^0 \in \mathbb{R}$, $i = 1, ..., m$ tais que

$$\alpha_0[(1-\lambda)\nabla g_0(0) + \lambda a_0] + \sum_{i=1}^{m} z_i^0[(1-\lambda)\nabla g_i(0) + \lambda a_i] = 0 \tag{3}$$

Isto implica

$$\alpha_0 a_0 + \sum_{i=1}^{m} z_i^0 a_i = 0 \tag{4}$$

Como a origem é um extremo de (P), temos que a equação (4) tem uma solução $\alpha_0 = \mu_0$, $z_i = y_i^0 \in \mathbb{R}_+ \backslash \{0\}$, donde segue que a origem é um extremo de (1), (2), para cada $\lambda \in [0, 1]$.

Agora, provaremos que o extremo $x = 0$ é uniformemente isolado com respeito a $\lambda \in [0, 1]$.

Caso contrário, existiria um $\lambda_0 \in [0, 1]$ para o qual o problema (1), (2) admite um extremo $x_0 \neq 0$, com norma arbitrariamente pequena. Neste caso, existem multiplicadores não simultaneamente nulos μ_1, $y_i^1 \geq 0$, $i = 1, ..., m$, tais que

$$\mu_1[(1-\lambda_0)\nabla g_0(x_0) + \lambda_0 a_0] + \sum_{i=1}^{m} y_i^1[(1-\lambda_0)\nabla g_i(x_0) + \lambda_0 a_i] = 0 \tag{5}$$

$$y_i^1[(1-\lambda_0)g_i(x_0) + \lambda_0 < a_i, x_0 >] = 0, \ i = 1, ..., m \tag{6}$$

Mas, para $||x_0||$ suficientemente pequena, os vetores $(1-\lambda_0)\nabla g_i(x_0) + \lambda_0 a_i$, $i = 1, ..., m$ são linearmente independentes e, portanto, se $\mu_1 = 0$, teríamos $y_i^1 = 0$ para todo i, o que é absurdo. Logo $\mu_1 > 0$.

Além disso, $y_i^1 > 0$, para todo i se $||x_0||$ for suficientemente pequena.

Assim, segue de (6) que

$$(1-\lambda_0)g_i(x_0) + \lambda_0 < a_i, x_0 >= 0, \ i = 1, ..., m \tag{7}$$

Como os vetores a_i são linearmente independentes, segue, do Teorema da Função Inversa, que $x_0 = 0$, absurdo.

Portanto, a origem é um extremo de (P), uniformemente isolado com respeito a λ.

Além disso, a origem é mínimo local de

$$\min < a_0, x >$$
$$\text{sujeito a: } < a_i, x > \leq 0, \; i = 1, ..., N$$

De fato, seja $\overline{x}$ um extremo deste último problema.

Neste caso, existem $\overline{\mu} \in \mathbb{R}^+, \overline{y} = (\overline{y}_1, ..., \overline{y}_N) \in \mathbb{R}_+^N$, não simultaneamente nulos, tais que

$$\overline{\mu} a_0 + \sum_{i=1}^{N} \overline{y}_i a_i = 0$$

Mas os vetores a_i são linearmente independentes, temos que $\overline{\mu} \neq 0$, e deste modo, podemos escrever

$$< a_0, \overline{x} >= - \sum_{i=1}^{N} \frac{\overline{y}_i}{\overline{\mu}} < a_i, \overline{x} > \geq 0 = < a_0, 0 >$$

Portanto, a origem é mínimo local.

Deste modo, a família de problema (1), (2) satisfazem as hipóteses do Teorema 4.5, e, portanto, a origem é ponto de mínimo local de (P). ■

4.5 As deformações lineares e um teorema de invariância do mínimo global

Sob certas hipóteses, vimos que é possível relacionar os mínimos locais do problema "referência " com os do problema "deformado ". A seguir, definimos um tipo de deformação que preserva a propriedade de um extremo ser mínimo global.

Antes, porém, definimos

Definição 4.11 *A família parametrizada de problemas*

$$(P(\lambda)) \begin{cases} \textit{minimizar } f(x;\lambda) = (1-\lambda) f_0(x) + \lambda f_1(x) \\ \textit{sujeito a: } g_i(x;\lambda) \equiv g_i(x) \leq 0, i = 1, ..., m; x \in \mathbb{R}^N \end{cases}$$

com $\lambda \in [0,1]$ é chamada ***deformação linear não singular*** *se satisfaz:*

1. *As funções $f(x;\lambda), \nabla_x f(x;\lambda)$ são contínuas em $\mathbb{R}^N \times [0,1]$ e $g_i^0(x), \nabla_x g_i(x)$ são contínuas em $\mathbb{R}^N$.*

2. *Para todo $\lambda \in [0,1]$ a origem é o único extremo do problema $(P(\lambda))$*

Teorema 4.12 *Suponha que a família de problemas $(P(\lambda))$ é uma deformação linear não singular e que o conjunto $\mathcal{D} = \{x \in \mathbb{R}^N : g_i(x) \leq 0$ para $i = 1, ..., m\}$ não é limitado. Além disso, suponha que para cada $x \in \mathcal{D}\backslash\{0\}$ exista um vetor $s(x) \in \mathbb{R}^N$ tal que $< \nabla f_0(x), s(x) >< 0$, $< \nabla g_i(x), s(x) >< 0$ para todo $i \in I(x)$ (onde $I(x)$ é o conjunto dos índices de restrições ativas em x) e que $\lim\limits_{\substack{x\in\mathcal{D}\\ |x|\to\infty}} f_0(x) = \infty$. Sob tais condições, a origem é mínimo global de (P(1)).*

Prova:

Fixe um $c > 0$ e considere o seguinte problema:

$$(P_c)\begin{cases} \min f_1(x) \\ g_i(x) \leq 0 \\ f_0(x) - c \leq 0 \end{cases}$$

com $x \in \mathbb{R}^N$, $i = 1, ..., m$.

Seja o conjunto $G = \{x \in \mathbb{R}^N : f_0(x) \leq c\}$.

Temos que $G \cap \mathcal{D}$ é compacto.

De fato, claramente temos que este conjunto é fechado pois as funções g_i, f_0 são contínuas em $\mathbb{R}^N$,e portanto,os conjuntos G e Q_0 são fechados, donde segue que $G \cap \mathcal{D}$ é fechado.

Além disso, suponha por absurdo que $G \cap \mathcal{D}$ não é limitado.

Neste caso, existe uma seqüência $(x_n) \subset G \cap \mathcal{D}$, com $|x_n| \to \infty$, quando $n \to \infty$ e, portanto, teríamos

$$f_0(x_n) \to \infty$$

contrariando $x_n \in G$ para cada n. Absurdo.

Portanto $G \cap \mathcal{D}$ é compacto.

Assim o problema (P_c) tem uma solução x_0.

Mostraremos que $x_0 = 0$.

De fato, segue do Teorema 2.3 que existem multiplicadores $\mu_0, \nu_0 \in \mathbb{R}_+, y_0 = (y_1^0, ..., y_m^0) \in \mathbb{R}_+^m$, não todos nulos tais que

$$\mu_0 \nabla f_1(x_0) + \nu_0 \nabla f_0(x_0) + \sum_{i=1}^{m} y_i^0 \nabla g_i(x_0) = 0 \tag{1}$$

$$y_i^0 g_i(x_0) = 0, i = 1, ..., m \tag{2}$$

$$\nu_0 (f_0(x_0) - c) = 0 \tag{3}$$

Multiplicando a equação (1) pelo vetor $s(x_0)$, obtemos:

$$\mu_0 < \nabla f_1(x_0), s(x_0) > + \nu_0 < \nabla f_0(x_0), s(x_0) > + \sum_{i=1}^{m} y_i^0 < \nabla g_i(x_0), s(x_0) >= 0 \tag{4}$$

Suponha que se tenha $\mu_0 + \nu_0 = 0$. Isto implicaria $\mu_0 = \nu_0 = 0$, e portanto, teríamos

$$\sum_{i=1}^{m} y_i^0 < \nabla g_i(x_0), s(x_0) >= 0$$

e como $< \nabla g_i(x_0), s(x_0) >< 0$ para todo i , teríamos $y_i^0 = \mu_0 = \nu_0 = 0$ para todo i, o que é absurdo. Logo $\mu_0 + \nu_0 > 0$.

Assim, a equação (1) pode ser rescrita como

$$\frac{\mu_0}{\mu_0 + \nu_0} \nabla f_1(x_0) + \frac{\nu_0}{\mu_0 + \nu_0} \nabla f_0(x_0) + \sum_{i=1}^{m} \frac{y_i^0}{\mu_0 + \nu_0} \nabla g_i(x_0) = 0$$

$$\frac{y_i^0}{\mu_0 + \nu_0} g_i(x_0) = 0, i = 1, ..., m$$

e, portanto, x_0 é extremo do problema

$$\begin{array}{l} \text{minimizar } \dfrac{\mu_0}{\mu_0 + \nu_0} f_1(x_0) + \dfrac{\nu_0}{\mu_0 + \nu_0} f_0(x_0) \\ \text{sujeito a: } g_i(x_0) \leq 0, i = 1, ..., m \end{array}$$

Portanto, segue pela definição de deformação linear não degenerada que $x_0 = 0$.

Desta forma, temos que para todo $c > 0$ a origem é mínimo global de (3)-(5) e, portanto, é mínimo global do problema $(P(1))$■

O Teorema 4.12 pode ser utilizado na prova de desigualdades. No próximo exemplo, provaremos a conhecida desigualdade das médias aritmética e geométrica.

Exemplo 4.13 *Sejam $x = (x_1, ..., x_N)$, $x_i \geq 0$. Considere:*

$$M(x) = \frac{1}{N}\sum_{i=1}^{N} x_i$$
$$\Gamma(x) = (\prod_{i=1}^{N} x_i)^{\frac{1}{N}}$$

Queremos provar que:

$$M(x) \geq \Gamma(x) \tag{1}$$

Faremos a mudança de variável $x_i = u_i^{2N}$ e consideraremos a família de funções:

$$f(u_1, ..., u_N; \lambda) = \frac{1}{N}\sum_{i=1}^{N} u_i^{2N} - \lambda(\prod_{i=1}^{N} u_i^2), \lambda \in [0, 1]. \tag{2}$$

Claramente, para $\lambda = 0$ temos que a origem é ponto de mínimo de $f(u_1, ..., u_N; \lambda)$.

Para provarmos que $f(u_1, ..., u_N; \lambda)$ é uma deformação não degenerada, basta mostrar que, para todo $\lambda \in [0, 1]$ o sistema

$$f'_{u_i}(u_1, ..., u_N; \lambda) = 0, i = 1, ..., N \tag{3}$$

admite apenas a solução nula.

Segue da equação (2) que (3) pode ser rescrita como

$$u_i^{2N-1} - \lambda u_i(\prod_{i \neq j} u_j^2) = 0, i = 1, ..., N \tag{4}$$

e, de (4) que a solução $u = (u_1, ..., u_N)$ da equação (3) é tal que $u_1 = ... = u_N$.

De fato, por (4), temos que

$$u_1^{2N-1} = \lambda u_1(\prod_{j \neq 1} u_j^2)$$
$$u_2^{2N-1} = \lambda u_2(\prod_{j \neq 2} u_j^2)$$

Note que se $\lambda = 0$*, o ponto* $u = (u_1, ..., u_N) = (0, ..., 0)$ *é a única solução do sistema (4).*

Das duas últimas equações, temos $u_1 = 0$ *se e somente se* $u_2 = 0$ *e, além disso, obtemos* $u_1^{2N} = u_2^{2N} = \lambda \prod_{i=1}^{N} u_j^2$*, donde segue que* $|u_1| = |u_2|$.

Suponha $\lambda \neq 0$.

Neste caso, temos que

$$\prod_{i=3}^{N} u_j^2 = \frac{u_2^{2(N-1)}}{\lambda u_1} = \frac{u_1^{2(N-1)}}{\lambda u_2}.$$

Portanto:

$$\lambda u_1^{2(N+1)+1} = \lambda u_2^{2(N+1)+1}$$

e, assim, u_1 *e* u_2 *têm o mesmo sinal. Logo,* $u_1 = u_2$.

Repetindo este raciocínio, obtemos $u_1 = ... = u_N$.

Seja $u_1 = ... = u_N = a$*. Substituindo este valor em (4), obtemos*

$$a^{2N-1}(1 - \lambda) = 0$$

para todo $\lambda \in [0, 1]$ *e portanto, temos* $a = 0$.

Assim, temos que $x(\lambda) \equiv 0$ *é mínimo local de* $f(u; 0)$ *e* $f(u; \lambda)$ *é uma deformação não degenerada. Portanto, o Teorema 4.12 garante que a origem é solução de* $f(u; 1)$ *e, assim, temos satisfeita a desigualdade (1).*

Parte II

O problema de tempo contínuo

O chamado problema de tempo contínuo foi introduzido em 1953 por Bellman [3], no estudo de certos problemas de produção industrial conhecidos por "bottleneck processes " e é formulado da seguinte maneira:

$$\left.\begin{array}{l} \text{minimizar } \phi(x) = \int_0^T f(t, x(t))dt \\ \text{sujeito a:} \\ g_i(t, x(t)) \leq 0 \text{ q.t.p. em } [0, T] \\ i \in I = \{1, ..., m\} \\ x \in X \end{array}\right\} \qquad \text{(CNP)}$$

onde X é um subconjunto não vazio, aberto e convexo do espaço $L_n^\infty[0, T]$ das funções definidas no intervalo $[0, T]$ a valores em $\mathbb{R}^n$, Lebesgue- mensuráveis, essencialmente limitadas. Este espaço, munido da norma $||\cdot||_\infty$ é um espaço de Banach, com esta norma definida por

$$||x||_\infty = \max_{j=1,...,n} ess \sup\{|x_j(t)|;\ t \in [0, T]\},\ x \in L_n^\infty[0, T]$$

Admitiremos ainda, que $f(t, x(t)) = \Gamma(x)(t)$ e $g(t, x(t)) = (g_1(t, x(t)), ..., g_m(t, x(t))) = \gamma(x)(t)$ onde $\gamma : X \to \Lambda_1^m[0, T]$ e $\Gamma : X \to \Lambda_1^1[0, T]$; o conjunto $\Lambda_1^m[0, T]$ é constituído das funções definidas em $[0, T]$ com valores em $\mathbb{R}^m$, Lebesgue mensuráveis, essencialmente limitadas e consideraremos o espaço $\Lambda_1^m[0, T]$ munido da norma $||\cdot||_1$, definida por

$$||y||_1 = \max_{j=1,...,m} \int_0^T |y_j(t)|dt,\ x \in \Lambda_1^m[0, T].$$

Observe que na formulação de (CNP) não foi feita nenhuma hipótese de diferenciabilidade sobre os funcionais envolvidos.

Em [26], Zalmai propõe condições necessárias de otimalidade para o caso diferenciável. O que faremos no Capítulo 5 é generalizar tais resultados para o caso não diferenciável, isto é, para o problema que formulamos acima.

Sob certas hipóteses de convexidade generalizada, Zalmai prova, ainda para o caso diferenciável, condições suficientes de otimalidade em [27]. No Capítulo 6, adaptaremos a noção de invexidade dada no Capítulo 3 ao contexto do problema de tempo contínuo e obteremos condições suficientes dos tipos Fritz John e Karush-Kuhn-Tucker para o problema (CNP).

Finalmente, no Capítulo 7 comentaremos as dificuldades encontradas na utilização do método de deformação ao problema (CNP).

Capítulo 5

Condições necessárias

Em [25], Zalmai prova uma generalização do Teorema de Gordan, através do qual demonstra uma condição necessária de otimalidade para o problema (CNP) (diferenciável) na forma de uma regra de multiplicadores. Procederemos de maneira análoga neste capítulo, cujas referências básicas são [4] e [5].

Iniciamos com o seguinte lema (generalização do Lema 2.2):

Lema 5.1 *(Teorema de Gordan Generalizado)*

Sejam A um subconjunto de $[0,T]$ com medida de Lebesgue positiva e X um subconjunto não vazio e convexo de $L_n^{\infty}[0,T]$ e as funções $p_i : A \times V \to \mathbb{R}$ onde V é um subconjunto aberto de $\mathbb{R}^n$, tal que $\{x(t) \in \mathbb{R}^n : g_i(t,x(t)) \leq 0$ q.t.p. em $[0,T], i \in I$ e $x \in X\} \subset V$ e $p_i(t,x(t)) = \pi_i(x)(t)$, $\pi = (\pi_1, ..., \pi_m) : X \to \Lambda_1^m(A)$ e π_i é convexo com respeito a x $(t \in A)$. Então, exatamente uma das afirmações seguintes é verdadeira:

1. *Existe um $x \in X$ tal que $p_i(t,x(t)) < 0$ q.t.p. $t \in A$ $(i \in I)$;*

2. *Existe uma função $u \in L_{\infty}^m(A) \backslash \{0\}$, $u_i(t) \geq 0$ q.t.p. $t \in A$ $(i \in I)$ tal que*

$$\int_A \sum_{i=1}^{m} u_i(t) p_i(t,x(t)) dt \geq 0$$

para cada $x \in X$.

A demonstração deste resultado segue as mesmas linhas da demonstração do Teorema 3.2 de [25], bastando substituir o intervalo $[0, T]$ pelo conjunto A e fazer os ajustes necessários.

Em nossa discussão posterior, assumimos em (CNP) que o conjunto $\{x \in X : g_i(t, x(t)) \leq 0$ q.t.p. em $[0, T], i \in I\}$ (conjunto das soluções factíveis para (CNP)) é não vazio e que existe V subconjunto aberto de $\mathbb{R}^n$ tal que

$$\{x(t) \in \mathbb{R}^n : g_i(t, x(t)) \leq 0 \text{ q.t.p. em } [0, T], i \in I \text{ e } x \in X\} \subset V$$

Assumiremos, ainda, que a função $t \mapsto f(t, x(t))$ é Lebesgue mensurável e integrável, para cada $x \in X$.

Sejam dados $\overline{x}$ factível para (CNP) e $i \in I$.

Definimos o conjunto

$$A_i(\overline{x}) = \{t : g_i(t, \overline{x}(t)) = 0\}.$$

Além disso, assumiremos que as funções $f(t, \cdot)$, $g_i(t, \cdot)$ são lipschitz próximo a todo $x \in V$ $(t \in [0, t])$. (Podemos assumir uma mesma constante, localmente, para todos os funcionais envolvidos).

Feitas estas considerações, podemos extender as noções de derivadas direcionais generalizadas ao contexto de tempo contínuo.

Sejam dados $\overline{x} \in X$ e $h \in L^n_\infty[0, T]$. Definimos:

$$g_i^0(t, \overline{x}(t); h(t)) := \gamma_i^0(\overline{x}; h)(t) = \limsup_{\substack{y \to \overline{x} \\ \lambda \to 0^+}} \frac{\gamma_i(y + \lambda h)(t) - \gamma_i(y)(t)}{\lambda}$$

$$f^0(t, \overline{x}(t); h(t)) := \Gamma^0(\overline{x}; h)(t) = \limsup_{\substack{y \to \overline{x} \\ \lambda \to 0^+}} \frac{\Gamma(y + \lambda h)(t) - \Gamma(y)(t)}{\lambda}.$$

Como as funções $f(t, \cdot)$ e $g_i(t, \cdot)$ $(i \in I)$ são Lipschitz próximo a qualquer $x \in V$ $(t \in [0, T])$ segue $g_i^0(t, \overline{x}(t); h(t))$ e $f^0(t, \overline{x}(t); h(t))$ são finitas q.t.p. em $[0, T]$.

Além disso, as funções $t \mapsto f^0(t, \overline{x}(t); h(t))$ e $t \mapsto g_i^0(t, \overline{x}(t); h(t))$ são Lebesgue mensuráveis e integráveis para cada $\overline{x} \in X$ e todo $h \in L^n_\infty[0, T]$.

Antes de prosseguirmos, definimos dois cones, em termos dos quais daremos uma caracterização geométrica de otimalidade local para (CNP):

$$K(\phi,\bar{x}) = \{h \in L^n_\infty[0,T] : \phi^0(\bar{x},h) < 0\}$$

$$K(g_i,\bar{x}) = \{h \in L^n_\infty[0,T] : g_i^0(t,\bar{x}(t);h(t)) < 0 \text{ q.t.p. } t \in A_i(\bar{x})\}$$

onde $i \in I$.

O teorema seguinte nos diz que se $\bar{x}$ é solução de (CNP), então a intersecção destes cones é vazia.

Teorema 5.2 *(caracterização geométrica de otimalidade) Se $\bar{x}$ é solução local de (CNP), então*

$$\cap_{i\in I} K(g_i,\bar{x}) \cap K(\phi,\bar{x}) = \phi. \tag{**}$$

Prova:

Suponhamos por absurdo que a intersecção (**) não seja vazia. Seja $h \in \cap_{i\in I} K(g_i,\bar{x}) \cap K(\phi,\bar{x})$.

Provaremos que existe um $\delta_0 > 0$ tal que

$$g_i(t,\bar{x}(t) + \lambda h(t)) \leq 0 \text{ q.t.p. } t \in [0,T]$$

$$\phi(\bar{x} + \lambda h) < \phi(\bar{x})$$

para todo $0 < \lambda < \delta_0$.

De fato, para cada $i \in I$, temos

$$g_i^0(t,\bar{x}(t);h(t)) = \limsup_{\substack{y\to\bar{x}\\ \lambda\downarrow 0}} \frac{\gamma_i(y+\lambda h)(t) - \gamma_i(y)(t)}{\lambda} < 0 \text{ q.t.p. } t \in A_i(\bar{x}), i \in I.$$

Em particular, para $y \equiv \bar{x}$, temos:

$$\limsup_{\lambda \downarrow 0} \frac{\gamma_i(\overline{x} + \lambda h)(t) - \gamma_i(\overline{x})(t)}{\lambda} < 0 \text{ q.t.p. } t \in A_i(\overline{x}), i \in I.$$

e, portanto,

$$\limsup_{\lambda \downarrow 0} \gamma_i(\overline{x} + \lambda h)(t) - \gamma_i(\overline{x})(t) < 0 \text{ q.t.p. } t \in A_i(\overline{x}), i \in I.$$

Como cada γ_i é lipschitz, temos que também é contínua.

Então, existe $\delta_i > 0$ tal que $\gamma_i(\overline{x} + \lambda h)(t) < 0$ para cada $\lambda \in (0, \delta_i)$ $(i \in I,\ t \in A_i(\overline{x}))$

Se $t \notin A_i(\overline{x})$, então, $g_i(t, \overline{x}(t)) < 0$ e como $g_i(t, \cdot)$ é contínua, temos que existe um $\lambda_0 > 0$ suficientemente pequeno, tal que

$$g_i(t, \overline{x}(t) + \lambda h(t)) < 0 \text{ para } \lambda \in (0, \lambda_0).$$

Seja $\delta' = \min_{i \in I}\{\delta_i, \lambda_0\}$. Então, temos que

$$\gamma_i(\overline{x} + \lambda h)(t) = g_i(t, \overline{x}(t) + \lambda h(t)) \leq 0 \text{ q.t.p. } t \in [0, T], i \in I$$

para todo $0 < \lambda < \delta'$.

Por outro lado:
$$\phi^0(\overline{x}; h) = \limsup_{\substack{y \to \overline{x} \\ \lambda \downarrow 0}} \frac{\phi(y + \lambda h) - \phi(y)}{\lambda} = \limsup_{\substack{y \to \overline{x} \\ \lambda \downarrow 0}} \frac{\int_0^T [f(t, (y + \lambda h)(t)) - f(t, y(t))]dt}{\lambda} < 0$$
Em particular, para $y \equiv \overline{x}$ obtemos,

$$\limsup_{\lambda \downarrow 0} \frac{\int_0^T [f(t, (\overline{x} + \lambda h)(t)) - f(t, \overline{x}(t))]dt}{\lambda} < 0$$

Como $\lambda > 0$, obtemos

$$\limsup_{\lambda \downarrow 0} \int_0^T [f(t, (\overline{x} + \lambda h)(t)) - f(t, \overline{x}(t))]dt < 0$$

Como $f(t, \cdot)$ é lipschitz, então é contínua, donde segue que existe um $\widehat{\delta} > 0$ tal que

$$\int_0^T f(t,(\overline{x}+\lambda h)(t))dt < \int_0^T f(t,\overline{x}(t))dt$$

para todo $\lambda \in (0,\widehat{\delta})$.

Ou seja

$$\phi(\overline{x}+\lambda h) < \phi(\overline{x})$$

para todo $\lambda \in (0,\widehat{\delta})$.

Seja $\delta_0 = \min\{\delta', \widehat{\delta}\}$.

Segue que

$$g_i(t,\overline{x}(t)+\lambda h(t)) \leq 0 \text{ q.t.p. } t \in [0,T], i \in I$$

e também

$$\phi(\overline{x}+\lambda h) < \phi(\overline{x})$$

para $\lambda \in (0,\delta_0)$, contrariando a otimalidade de $\overline{x}$. ■

Com o auxílio do Lema 5.1 obtém-se uma condição necessária de otimalidade na forma da seguinte regra de multiplicadores:

Teorema 5.3 *(Fritz John)*

Sejam $\overline{x}$ factível para (CNP) e suponha que os funcionais $f(t,\cdot)$, $g_i(t,\cdot)$ Lipschitz próximo a $\overline{x}(t)$ e suponha que $\overline{x}$ seja solução local de (CNP). Então existem $\overline{u}_0 \in \mathbb{R}$, $\overline{u}_i \in L^m_\infty[0,T]$, $i \in I$ tais que

$$0 \in \int_0^T [\overline{u}_0 \partial_x f(t,\overline{x}(t)) + \sum_{i\in I} \overline{u}_i(t)\partial_x g_i(t,\overline{x}(t))]dt \tag{1}$$

$$\overline{u}_0 \geq 0, \ \overline{u}(t) \geq 0 \text{ q.t.p. em } [0,T] \tag{2}$$

$$(\overline{u}_0, \overline{u}(t)) \not\equiv 0 \text{ q.t.p. em } [0,T] \tag{3}$$

$$\overline{u}_i(t) g_i(t, \overline{x}(t)) = 0 \text{ q.t.p. em } [0,T] \tag{4}$$

onde $\overline{u}(t) = (\overline{u}_1(t), ..., \overline{u}_m(t))$.

Prova:

Primeiramente, suporemos que o problema (CNP) possui apenas uma restrição de desigualdade, digamos, $g(t, x(t)) \leq 0$ q.t.p. em $[0,T]$.

Consideraremos, então, os conjuntos:

$$A(\overline{x}) = \{t \in [0,T] : g(t, \overline{x}(t)) = 0\}$$

$$K(g, \overline{x}) = \{h \in L^n_\infty[0,T] : g^0(t, \overline{x}(t); h(t)) < 0, t \in A(\overline{x})\}.$$

Sejam $\overline{x}$ factível para (CNP) uma solução local.

Provaremos que existem $\overline{u}_0 \in \mathbb{R}$, $\overline{u} \in L^n_\infty[0,T]$ tais que

$$0 \leq \int_0^T [\overline{u_0} f^0(t, \overline{x}(t); h(t)) + \overline{u}(t) g^0(t, \overline{x}(t); h(t))] dt \text{ para todo } h \in L^n_\infty[0,T]; \tag{5}$$

$$\overline{u_0} \geq 0, \overline{u}(t) \geq 0 \text{ q.t.p. } t \in [0,T] \tag{6}$$

$$(\overline{u_0}, \overline{u}(t)) \neq 0 \text{ q.t.p. } t \in [0,T] \tag{7}$$

Como $\overline{x}$ é solução de (CNP) segue do Teorema 5.2 que $K(g, \overline{x}) \cap K(\phi, \overline{x}) = \phi$. Portanto, não existe $h \in L^n_\infty[0,T]$ tal que $\phi^0(\overline{x}; h) < 0$ e $g^0(t, \overline{x}(t); h(t)) < 0$ q.t.p. $t \in A(\overline{x})$.

Assim, segue do Lema 5.1 (Teorema de Gordan Generalizado) que existem u_0, $u \in L^\infty[0,T]$, $u(t) \geq 0$ q.t.p. $t \in A(\overline{x})$ tais que

$$0 \leq \int_{A(\overline{x})} [u_0(t)\phi^0(\overline{x};h) + u(t)g^0(t,\overline{x}(t);h(t))]dt \text{ para todo } h \in L^n_\infty[0,T]. \tag{8}$$

Sejam $\overline{u}_0 = \int_{A(\overline{x})} u_0(t)dt$ e $\overline{u}(t) = \begin{cases} u(t) \text{ para } t \in A(\overline{x}); \\ 0 \text{ caso contrário} \end{cases}$.

Portanto:

$$0 \leq \overline{u}_0\phi^0(\overline{x};h) + \int_0^T \overline{u}(t)g^0(t,\overline{x}(t);h(t))dt \leq \int_0^T [u_0 f^0(t,\overline{x}(t);h(t)) + \overline{u}(t)g^0(t,\overline{x}(t);h(t))]dt$$

onde a a primeira desigualdade segue de (8) e da definição de $\overline{u}_0$ e de $\overline{u}$ e a segunda desigualdade segue do Lema de Fatou (v. Royden, [21]). Claramente, $\overline{u}_0$ e $\overline{u}$ satisfazem (5)-(7) .

Logo:

$$0 \leq \int_0^T [\overline{u}_0\sigma_{\partial_x f(t,\overline{x}(t))}(h(t)) + \overline{u}(t)\sigma_{\partial_x g(t,\overline{x}(t))}(h(t))]dt = \int_0^T \sigma_{[\overline{u}_0\partial_x f(t,\overline{x}(t))+\overline{u}(t)\partial_x g(t,\overline{x}(t))]}(h(t))dt$$

para todo $h \in L^n_\infty[0,T]$,onde a primeira desigualdade segue de (5) e a igualdade segue das propriedades da integral de um funcional suporte (vistas no Capítulo 1 deste trabalho)

Em particular, a desigualdade anterior segue válida para a função $h(t) \equiv v$ ($t \in [0,T]$) e v é um vetor arbitrário de $\mathbb{R}^n$.

Por outro lado, a multifunção $t \mapsto \overline{u}_0\partial_x f(t,\overline{x}(t)) + \overline{u}(t)\partial_x g(t,\overline{x}(t))$ assume valores compactos em $\mathbb{R}^n$ e é integravelmente limitada.

Então:

$$0 \leq \int_0^T \sigma_{[\overline{u}_0\partial_x f(t,\overline{x}(t))+\overline{u}(t)\partial_x g(t,\overline{x}(t))]}(v)dt = \sigma_{\int_0^T [\overline{u}_0\partial_x f(t,\overline{x}(t))+\overline{u}(t)\partial_x g(t,\overline{x}(t))]dt}(v)$$

para todo $v \in \mathbb{R}^n$, e, novamente, pelo Lema de Hörmander , obtemos,

$$0 \in \int_0^T [\overline{u}_0\partial_x f(t,\overline{x}(t)) + \overline{u}(t)\partial_x g(t,\overline{x}(t))]dt$$

o que conclui a prova do teorema para o caso $m = 1$.

Agora, provaremos o caso em que (CNP) possui $m \geq 2$ restrições de desigualdades e seja $\overline{x}$ uma solução local de (CNP).

Seja $g(t,x(t)) = \max_{i=1,\dots,m} g_i(t,x(t))$ q.t.p. em $[0,T]$.

Claramente, $\overline{x}$ é uma solução local de

$$\begin{cases} \text{minimizar } \phi(x) \\ \text{sujeito a:} \\ g(t,x(t)) \leq 0 \text{ q.t.p. em } t \in [0,T] \\ x \in X \end{cases}$$

Para cada $t \in [0,T]$ e cada $x \in X$ definimos: $I(t,x) = \{i : g_i(t,x(t)) = g(t,x(t))\}$.

Aplicando o resultado anterior a este problema, temos que existem $\overline{u}_0 \in \mathbb{R}$, $u \in L^\infty[0,T]$ satisfazendo (5)-(7) e tal que

$$0 \in \int_0^T [\overline{u}_0 \partial_x f(t,\overline{x}(t)) + u(t)\partial_x g(t,\overline{x}(t))]dt \tag{9}$$

Da definição de integral de multifunções, segue que existe uma função mensurável e tal que $e(t) \in \partial_x g(t,\overline{x}(t))$ q.t.p. $t \in [0,T]$ tal que

$$0 \in \int_0^T [\overline{u}_0 \partial_x f(t,\overline{x}(t)) + u(t)e(t)]dt \tag{10}$$

Provaremos agora o seguinte lema:

Lema 5.4 *Existe uma função $v \in L^\infty_m[0,T]$, $v(t) \geq 0$ q.t.p. em $t \in [0,T]$ tal que*

- $v_i(t) = 0$ *quando* $g_i(t,\overline{x}(t)) \neq g(t,\overline{x}(t))$, $i = 1,\dots,m$;
- $\sum_{i=1}^m v_i(t) = 1$ *q.t.p.* $t \in [0,T]$;
- $e(t) \in \sum_{i=1}^m v_i(t)\partial_x g(t,\overline{x}(t))$.

Prova: (do Lema 5.4)

De fato, pode-se provar (v. [13], p.47) que para cada $t \in [0,T]$, temos

$$\partial_x g(t,\overline{x}(t)) \subset \mathrm{co}\{\partial_x g_i(t,\overline{x}(t)) : i \in I(t,\overline{x})\}$$

Como $e(t) \in \partial_x g(t,\overline{x}(t))$ q.t.p. $t \in [0,T]$ temos

$$e(t) \in \text{co}\{\partial_x g_i(t, \overline{x}(t)) : i \in I(t, \overline{x})\} \text{ q.t.p. } t \in [0, T].$$

Agora, para cada t, defina o conjunto

$V(t) = \{v = (v_1, ..., v_m) \in \mathbb{R}^m : \sum_{i=1}^m v_i = 1, v_i \geq 0, v_i = 0$ se $g_i(t, \overline{x}(t)) < g(t, \overline{x}(t))$ e $e(t) \in \sum_{i=1}^m v_i \partial_x g_i(t, \overline{x}(t))\}$.

O conjunto $V(t)$ é fechado para cada t. Além disso, a multifunção $V(t)$ é não vazia para cada t, definida q.t.p. em $[0, T]$ e mensurável.

Pelo Teorema de Seleção Mensurável (v, [1], p. 90-91), existem funções mensuráveis

$$v_1, ..., v_m : [0, T] \to \mathbb{R},$$

tais que

$$v(t) = (v_1(t), ..., v_m(t)) \in V(t) \text{ q.t.p.}$$

o que prova o lema. ■

Retornemos à prova do Teorema 5.3.

Defina: $\overline{u}_i(t) = u(t)v_i(t)$.

Temos então:

$$0 \in \int_0^T [\overline{u}_0 \partial_x f(t, \overline{x}(t)) + u(t)e(t)]dt \subset \int_0^T [\overline{u}_0 \partial_x f(t, \overline{x}(t)) + \sum_{i=1}^m \underbrace{u(t)v_i(t)}_{=\overline{u}_i(t)} \partial_x g_i(t, \overline{x}(t))]dt$$

Mas

$$\int_0^T [\overline{u}_0 \partial_x f(t, \overline{x}(t)) + \sum_{i=1}^m u(t)v_i(t)\partial_x g_i(t, \overline{x}(t))]dt = \int_0^T [\overline{u}_0 \partial_x f(t, \overline{x}(t)) + \sum_{i=1}^m \overline{u}_i(t)\partial_x g_i(t, \overline{x}(t))]dt$$

o que conclui a prova do teorema. ■

Suponha no problema (CNP) que as funções $f(t, \cdot), g_i(t, \cdot)$ são Clarke regulares e defina a função Lagrangeana $L : L_n^\infty[0, T] \times \mathbb{R} \times L_m^\infty[0, T] \to \mathbb{R}$, onde

$$L(\overline{x}, \overline{u}_0, \overline{u}) = \int_0^T [\overline{u}_0 f(t, \overline{x}(t)) + \sum_{i=1}^m \overline{u}_i(t) g_i(t, \overline{x}(t))] dt.$$

A condição (2) do Teorema 5.3 pode ser rescrita como $0 \in \partial_x L(\overline{x}, \overline{u}_0, \overline{u})$.

De fato, se as funções forem regulares temos que

$$\partial(f + \sum_{i=1}^m g_i)(\overline{x}) = \partial f(\overline{x}) + \sum_{i=1}^m \partial g_i(\overline{x}).$$

Provaremos que se f é regular, então

$$\int_0^T \partial_x f(t, \overline{x}(t)) dt = \partial_x \int_0^T f(t, \overline{x}(t)) dt.$$

Primeiramente, notemos que se f é regular, então ϕ é regular.
De fato, temos $\limsup\limits_{\substack{\lambda \downarrow 0 \\ y \to \overline{x}}} \dfrac{f(t, (y + \lambda h)(t)) - f(t, y(t))}{\lambda} = \lim\limits_{\lambda \downarrow 0} \dfrac{f(t, (\overline{x} + \lambda h)(t)) - f(t, \overline{x}(t))}{\lambda}$.
Mas:

$$\begin{aligned}
\phi^0(\overline{x}; v) &= \limsup_{\substack{\lambda \downarrow 0 \\ y \to \overline{x}}} \int_0^T \frac{f(t, (y + \lambda h)(t)) - f(t, y(t))}{\lambda} dt \\
&\leq \int_0^T \limsup_{\substack{\lambda \downarrow 0 \\ y \to \overline{x}}} \frac{f(t, (y + \lambda h)(t)) - f(t, y(t))}{\lambda} dt \\
&= \int_0^T \lim_{\lambda \downarrow 0} \frac{f(t, (\overline{x} + \lambda h)(t)) - f(t, \overline{x}(t))}{\lambda} dt \\
&= \lim_{\lambda \downarrow 0} \int_0^T \frac{f(t, (\overline{x} + \lambda h)(t)) - f(t, \overline{x}(t))}{\lambda} dt \\
&= \phi'(\overline{x}; h)
\end{aligned}$$

e, portanto, ϕ é regular em $\overline{x}$.

Por outro lado,

$$
\begin{aligned}
\int_0^T f^0(t,\overline{x}(t);h(t))dt &= \int_0^T \lim_{\lambda\downarrow 0}\frac{f(t,(\overline{x}+\lambda h)(t))-f(t,\overline{x})}{\lambda}dt \\
&= \lim_{\lambda\downarrow 0}\frac{\phi(\overline{x}+\lambda h)-\phi(\overline{x})}{\lambda} \\
&= \phi'(\overline{x};h)=\phi^0(\overline{x};h)
\end{aligned}
$$

Logo, $\int_0^T \sigma_{\partial_x f(t,\overline{x}(t))}(h(t))dt = \sigma_{\partial_x\phi(\overline{x})}(h)$ para todo $h\in L_\infty^n[0,T]$.
Em particular, para $h(t)\equiv v\in\mathbb{R}^n$, obtemos

$$
\int_0^T \sigma_{\partial_x f(t,\overline{x}(t))}(v)dt = \sigma_{\partial_x\phi(\overline{x})}(v)
$$

para todo $v\in\mathbb{R}^n$.

Mas a multifunção $t\mapsto\partial_x f(t,\overline{x}(t))$ é integravelmente limitada e assume valores compactos. Neste caso,

$$
\int_0^T f^0(t,\overline{x}(t);v)dt = \int_0^T \sigma_{\partial_x f(t,\overline{x}(t))}(v)dt = \sigma_{\int_0^T \partial_x f(t,\overline{x}(t))dt}(v) = \sigma_{\partial\phi(\overline{x})}(v)
$$

para todo $v\in\mathbb{R}^n$.

Logo,

$$
\partial_x\phi(\overline{x}) = \int_0^T \partial_x f(t,\overline{x}(t))dt = \partial_x\int_0^T f(t,\overline{x}(t))dt.
$$

Assim,

$$
\begin{aligned}
\int_0^T [\overline{u_0}\partial_x f(t,\overline{x}(t)) + \sum_{i=1}^m \overline{u_i}\partial_x g_i(t,\overline{x}(t))]dt &= \partial_x[\int_0^T \overline{u_0} f(t,\overline{x}(t))dt] + \partial_x[\int_0^T\sum_{i=1}^m \overline{u_i} g_i(t,\overline{x}(t))dt] \\
&= \partial_x[\int_0^T \overline{u_0} f(t,\overline{x}(t)) + \sum_{i=1}^m \overline{u_i} g_i(t,\overline{x}(t))dt] \\
&= \partial_x L(\overline{x},\overline{u_0},\overline{u})
\end{aligned}
$$

Veremos a seguir que sob uma hipótese de regularidade sobre as restrições, podemos assumir no teorema anterior que o multiplicador associado ao funcional objetivo é diferente de zero.

Teorema 5.5 *(Karush-Kuhn-Tucker)*

Suponha que $\overline{x}$ seja uma solução local para (CNP) e que

$$\cap_{i\in I} K(g_i, \overline{x}) \neq \phi \tag{1}$$

Então, existem $\widetilde{u}_i \in L^\infty[0,T]$ $(i \in I)$ tais que

$$0 \in \int_0^T [\partial_x f(t, \overline{x}(t)) + \sum_{i\in I} \widetilde{u}_i(t) \partial_x g_i(t, \overline{x}(t))] dt \tag{2}$$

$$\widetilde{u}_i(t) \geq 0 \text{ q.t.p. } t \in [0,T], \ i \in I \tag{3}$$

$$\widetilde{u}_i(t) g_i(t, \overline{x}(t)) = 0 \text{ q.t.p. } t \in [0,T], \ i \in I \tag{4}$$

Prova:

Primeiramente, discutiremos o caso em que o problema (CNP) possui apenas uma restrição de desigualdade, digamos, $g(t, x(t)) \leq 0$ q.t.p. $t \in [0,T]$.

Pelo teorema anterior, temos que existem $\overline{u}_0 \in \mathbb{R}, \overline{u} \in L^\infty_n[0,T]$ tais que

$$0 \in \int_0^T [\overline{u}_0 \partial_x f(t, \overline{x}(t)) + \overline{u}(t) \partial_x g(t, \overline{x}(t))] dt$$

$$\overline{u}_0 \geq 0, \ \overline{u}(t) \geq 0 \text{ q.t.p. em } [0,T]$$

$$(\overline{u}_0, \overline{u}(t)) \not\equiv 0 \text{ q.t.p. em } [0,T]$$

$$\overline{u}_i(t) g(t, \overline{x}(t)) = 0 \text{ q.t.p. em } [0,T]$$

Suponha por absurdo, que $\overline{u}_0 = 0$.

Neste caso, temos,

$$0 \leq \int_0^T \overline{u}(t) g^0(t, \overline{x}(t); h(t)) dt \tag{5}$$

para todo $h \in L_n^\infty[0,T]$.

Assim, pelo Lema 5.1 (Teorema de Gordan Generalizado), não existe $h \in X$ tal que $g^0(t, \overline{x}(t); h(t)) < 0$ q.t.p. em $[0,T]$, o que contraria a condição de regularidade (1). Absurdo!

Façamos, então, $\widetilde{u}(t) = \dfrac{\overline{u}(t)}{\overline{u}_0}$.

Temos, então, $0 \leq \int_0^T [f^0(t, \overline{x}(t); h(t)) + \widetilde{u}(t) g^0(t, \overline{x}(t))] dt$ para todo $h \in L_n^\infty[0,T]$, e, portanto,

$$0 \in \int_0^T [\partial_x f(t, \overline{x}(t)) + \widetilde{u}(t) \partial_x g(t, \overline{x}(t))] dt$$

Agora, retornemos ao caso geral em que (CNP) possui $m \geq 2$ restrições de desigualdade.

Neste caso o problema (CNP) é equivalente a um problema com uma única restrição de desigualdade, $g(t, x(t)) \leq 0$ q.t.p. $t \in [0,T]$, onde $g(t, x(t)) = \max\limits_{i=1,\dots,m} g_i(t, x(t))$.

Provaremos a seguir que a condição de regularidade (1) implica $K(g, \overline{x}) \neq \phi$.

De fato, seja o conjunto

$$\Gamma(t) = \{v = (v_1, \dots, v_n) : v \in \partial_x g(t, \overline{x}(t)) \text{ e } g^0(t, \overline{x}(t); h(t)) = \sum_{i=1}^n v_i h_i(t),\ \forall h \in L_\infty^n[0,T]\}$$

O conjunto $\Gamma(t)$ é fechado e não vazio e Γ é uma multifunção mensurável.

Pelo Teorema da Seleção Mensurável [13], existe uma $\xi \in L_\infty^n[0,T]$ tal que $\xi(t) \in \Gamma(t)$ q.t.p. $t \in [0,T]$, ou seja,

$$\xi(t) \in \partial_x g(t, \overline{x}(t))$$

$$g^0(t, \overline{x}(t); h(t)) = \sum_{i=1}^n \xi_i(t) h_i(t) \text{ q.t.p. } t \in [0,T] \text{ e todo } h \in L_\infty^n[0,T]$$

Fixemos $h \in L_\infty^n[0,T]$. Aplicando o Lema 5.4 e o Lema de Hörmander, concluímos que existe $u(t) \in V(t)$ q.t.p. $t \in [0,T]$ (onde $V(t)$ é definido no Lema 5.4), tal que

$$\begin{aligned}
\xi(t) &\in \sum_{i\in I} u_i(t)\partial_x g_i(t,\overline{x}(t));\\
g^0(t,\overline{x}(t);h(t)) &\leq \sigma_{\sum_{i\in I} u_i(t)\partial_x g_i(t,\overline{x}(t))}(h(t))
\end{aligned}$$

q.t.p. $t \in [0,T]$.

Da desigualdade acima e do Lema de Hörmander, segue que

$$\begin{aligned}
g^0(t,\overline{x}(t);h(t)) &\leq \sum_{i\in I} u_i(t)\sigma_{\partial_x g_i(t,\overline{x}(t))} \text{ q.t.p. } t \in [0,T]\\
&= \sum_{i\in I} u_i(t) g_i^0(t,\overline{x}(t);h(t)) \text{ q.t.p. } t \in [0,T].
\end{aligned}$$

E, portanto,

$$\cap_{i\in I} K(g_i,\overline{x}) \neq \phi \text{ implica } K(g,\overline{x}) \neq \phi$$

Logo, segue do fato que provamos agora e do Teorema de Kuhn Tucker com uma restrição de desigualdade, o qual provamos anteriormente, que existe $u \in L_\infty[0,T]$ tal que

$$0 \in \int_0^T [\partial_x f(t,\overline{x}(t)) + u(t)\partial_x g(t,\overline{x}(t))]dt$$

$$u(t) \geq 0 \text{ q.t.p. } t \in [0,T]$$

$$u(t)g(t,\overline{x}(t)) = 0$$

e, usando a mesma argumentação da prova do Teorema 5.3 segue o resultado desejado. ■

Particularizando os resultados acima obtidos para o contexto diferenciável, obtemos os seguintes corolários:

Corolário 5.6 *(Fritz-John) Seja $\overline{x}$ factível para (CNP) e suponha que as funções f, g_i $(i \in I)$ sejam continuamente diferenciáveis em seu segundo argumento em $\overline{x}(t)$ sobre $[0,T]$ para cada $i \in I$, que as g_i sejam contínuas am seu segundo argumento nos pontos da imagem de $\overline{x}(t)$ em $[0,T]$ e que as funções $t \mapsto < \nabla f(t,\overline{x}(t)), h(t) >$ e $t \mapsto < \nabla g_i(t,\overline{x}(t)), h(t) >$ $(i \in I)$ sejam Lebesgue integráveis em $[0,T]$ para todo $h \in L_n^\infty[0,T]$. Se $\overline{x}$ é solução local de (CNP), então existem $\overline{u}_0 \in \mathbb{R}$, $\overline{u}_i \in L_\infty^m[0,T]$, $i \in I$ tais que*

$$\int_0^T < [\overline{u}_0 \nabla f(t,\overline{x}(t)) + \sum_{i=1}^m \overline{u}_i(t) \nabla g_i(t,\overline{x}(t))], h(t) > dt = 0$$

para todo $h \in L_n^\infty[0,T]$,

$$\overline{u}_0 \geq 0, \overline{u}_i(t) \geq 0 \text{ q.t.p. em } [0,T]$$

$$(\overline{u}_0, \overline{u}(t)) \neq 0 \text{ q.t.p. em } [0,T]$$

$$\overline{u}_i(t) g_i(t,\overline{x}(t)) = 0 \text{ q.t.p. em } [0,T],\ i \in I$$

(onde $\overline{u}(t) = (\overline{u}_1(t), ..., \overline{u}_m(t))$)

Notemos que sob as hipótese do corolário anterior, temos que

$$K(g_i,\overline{x}) = \{h \in L_\infty^n[0,T] : g_i'(t,\overline{x}(t); h(t)) < 0 \text{ q.t.p. } t \in A_i(\overline{x})\} \qquad (**)$$

e, portanto, pelo Teorema 5.5 segue que podemos assumir no Corolário 5.6, $\overline{u}_0 > 0$.

Destas considerações segue o seguinte

Corolário 5.7 *(Karush-Kuhn-Tucker) Suponha as hipóteses do corolário anterior e a condição de regularidade: $\bigcap_{i\in I} K(g_i,\overline{x}) \neq \phi$ onde cada $K(g_i,\overline{x})$ é definido por (**). Se $\overline{x}$ é solução local de (CNP), então,existem $\overline{u}_i \in L_\infty^m[0,T]$, $i \in I$ tais que*

$$\int_0^T < [\nabla f(t,\overline{x}(t)) + \sum_{i=1}^{m} \overline{u}_i(t) \nabla g_i(t,\overline{x}(t))], h(t) > dt = 0$$

para todo $h \in L_n^\infty[0,T]$,

$$\overline{u}_i(t) \geq 0 \text{ q.t.p. em } [0,T]$$

$$\overline{u}(t) \neq 0 \text{ q.t.p. em } [0,T]$$

$$\overline{u}_i(t) g_i(t,\overline{x}(t)) = 0 \text{ q.t.p. em } [0,T], \ i \in I$$

(onde $\overline{u}(t) = (\overline{u}_1(t), ..., \overline{u}_m(t))$)

Capítulo 6

Condições suficientes

Vimos no Capítulo 5 condições necessárias de otimalidade para o problema de tempo contínuo sob a forma de uma regra de multiplicadores. Veremos que, sob certas hipóteses de convexidade generalizada (neste caso, invexidade) é possível obter condições suficientes de otimalidade para (CNP) análogas às vistas no Capítulo 3. Para isto, deveremos adaptar a noção de invexidade dada no Capítulo 2 para o contexto de tempo contínuo.

Assim, neste capítulo, apresentaremos uma noção de invexidade adequada ao problema de tempo contínuo, e veremos que sob hipóteses de invexidade, as condições de Fritz John e de Karush-Kuhn-Tucker vistas na seção anterior resultam suficientes.

A referência para este capítulo .

Em nossa discussão, assumiremos que existe V aberto e convexo tal que

$$V \supset \{x(t) \in \mathbb{R}^n : x \in \mathbf{F}, t \in [0,T]\},$$

que $f, g_i : [0,T] \times V \longrightarrow \mathbb{R}$, que a aplicação $t \longmapsto f(t, x(t))$ é Lebesgue mensurável e integrável pata cada $x \in X$ e que as funções $f(t,\cdot)$ e $g_i(t,\cdot)$, $i \in I$ são localmente Lipschitz sobre V em $[0,T]$. Naturalmente, podemos assumir que a constante de Lipschitz é a mesma para todas as funções.

Admitiremos ainda, que $f(t,x(t)) = \Gamma(x)(t)$ e $g(t,x(t)) = (g_1(t,x(t)),...,g_m(t,x(t))) = \gamma(x)(t)$ onde $\gamma : X \rightarrow \Lambda_1^m[0,T]$ e $\Gamma : X \rightarrow \Lambda_1^1[0,T]$.

O conceito de invexidade visto na Capítulo 3 admite a seguinte (e natural) generalização:

Sejam U um subconjunto não vazio de Z e $\psi : U \to \mathbb{R}$ uma função localmente lipschitz. Dizemos que ψ é **invexa** em $\overline{z} \in U$ se existe uma função $\eta : U \times U \to Z$ tal que

$$\psi(z) - \psi(\overline{z}) \geq \psi^0(\overline{z}; \eta(z, \overline{z})).$$

Agora, adaptaremos esta última noção ao contexto tempo contínuo:

Definição 6.1 *Sejam U um subconjunto não vazio de $\mathbb{R}^n$, $\overline{x} \in X$ e $\psi : [0,T] \times U \to \mathbb{R}$ uma função localmente lipschitz em $[0,T]$. Dizemos que $\psi(t, \cdot)$ é* ***invexa*** *em $\overline{x}(t)$ respeito a U se existe uma função $\eta : U \times U \to \mathbb{R}^n$ tal que $t \mapsto \eta(x(t), \overline{x}(t))$ está em $L_\infty^n[0,T]$ e também*

$$\psi(t, x(t)) - \psi(t, \overline{x}(t)) \geq \psi^0(t, \overline{x}(t); \eta(x(t), \overline{x}(t))) \text{ q.t.p.em } [0,T] \text{ e todo } x \in X.$$

Além disso, dizemos que ψ é ***estritamente invexa*** *se é invexa e se a desigualdade acima é estrita para todo $x \neq \overline{x}$.*

Em [19], os autores obtém condições suficientes de otimalidade global para (CNP) do tipo Karush-Kuhn-Tucker e Fritz John sob hipóteses de invexidade. São os Teoremas 6.2 e 6.3, que a seguir enunciamos e demonstramos:

Teorema 6.2 *(Fritz John) Em (CNP), suponha $\overline{x}$ factível para (CNP), que $f(t, \cdot)$ é invexa em $\overline{x}(t)$ (com respeito a V) em $[0,T]$ e que para cada $i \in I$, $g_i(t, \cdot)$ é estritamente invexa em $\overline{x}(t)$ (com respeito a V) em $[0,T]$, com mesma $\eta(x(t), \overline{x}(t))$ para todas as funções. Suponha, ainda, que existam multiplicadores $\overline{\lambda_0} \in \mathbb{R}$, $\overline{\lambda} = (\overline{\lambda}_1, ..., \overline{\lambda}_m) \in L_\infty^m[0,T]$ tais que*

$$0 \leq \int_0^T [\overline{\lambda_0} f^0(t, \overline{x}(t); h(t)) + \sum_{i=1}^m \overline{\lambda}_i(t) g_i^0(t, \overline{x}(t); h(t))] dt \text{ para todo } h \in L_\infty^n[0,T] \tag{1}$$

$$\overline{\lambda_0} \geq 0,\ \overline{\lambda}(t) \geq 0 \text{ q.t.p. em } [0,T] \tag{2}$$

$$(\overline{\lambda_0}, \overline{\lambda}(t)) \neq 0 \text{ q.t.p. em } [0,T] \tag{3}$$

$$\overline{\lambda}_i(t) g_i(t, \overline{x}(t)) = 0 \text{ q.t.p. em } [0,T] \text{ para todo } i \in I \tag{4}$$

Então $\overline{x}$ é solução global de (CNP).

Prova:

Suponhamos por absurdo que $\overline{x}$ não seja solução de (CNP).

Neste caso, existe um $\widetilde{x}$ factível e distinto de $\overline{x}$, satisfazendo

$$\int_0^T f(t, \widetilde{x}(t))dt < \int_0^T f(t, \overline{x}(t))dt$$

Das hipóteses de invexidade e de invexidade estrita dos funcionais, segue que

$$f(t, \widetilde{x}(t)) - f(t, \overline{x}(t)) \geq f^0(t, \overline{x}(t); \eta(\widetilde{x}(t), \overline{x}(t))) \text{ q.t.p. em } [0,T]$$
$$g_i(t, \widetilde{x}(t)) - g_i(t, \overline{x}(t)) > g_i^0(t, \overline{x}(t); \eta(\widetilde{x}(t), \overline{x}(t))) \text{ q.t.p. em } [0,T],\ i \in I.$$

Mas $\widetilde{x}$ é factível e como $\overline{\lambda}_i(t) \geq 0$ q.t.p. $t \in [0,T]$, $i \in I$, segue

$$\overline{\lambda}_i(t) g_i(t, \widetilde{x}(t)) \leq 0 \text{ q.t.p. em } [0,T] \text{ e } i \in I.$$

Portanto, segue das desigualdades anteriores e de (2)-(4) que

$$0 > \int_0^T [\overline{\lambda_0} f^0(t, \overline{x}(t); \eta(\widetilde{x}(t), \overline{x}(t))) + \sum_{i=1}^m \overline{\lambda}_i(t) g_i^0(t, \overline{x}(t); \eta(\widetilde{x}(t), \overline{x}(t)))]dt$$

Mas $h(\cdot) = \eta(\widetilde{x}(\cdot), \overline{x}(\cdot)) \in L^n_\infty[0,T]$, contrariando (1).

Portanto, $\overline{x}$ é solução global de (CNP). ■

A seguir, veremos condições de otimalidade do tipo Karush-Kuhn-Tucker:

Teorema 6.3 *(Karush-Kuhn-Tucker) Em (CNP), suponha $\overline{x}$ factível para (CNP), que $f(t,\cdot)$ e $g_i(t,\cdot)$ são invexas em $\overline{x}(t)$ (com respeito a V) em $[0,T]$ e que para cada $i \in I$, com mesma $\eta(x(t), \overline{x}(t))$ para todas as funções. Suponha, ainda, que existam multiplicadores $\overline{\lambda} = (\overline{\lambda}_1, ..., \overline{\lambda}_m) \in L^m_\infty[0,T]$ tais que*

$$0 \leq \int_0^T [f^0(t, \overline{x}(t); h(t)) + \sum_{i=1}^m \overline{\lambda}_i(t) g_i^0(t, \overline{x}(t); h(t))] dt \text{ para todo } h \in L_\infty^n[0,T] \tag{1}$$

$$\overline{\lambda_i}(t) \geq 0 \text{ q.t.p. em } [0,T],\ i \in I \tag{2}$$

$$\overline{\lambda}_i(t) g_i(t, \overline{x}(t)) = 0 \text{ q.t.p. em } [0,T] \text{ para todo } i \in I \tag{3}$$

Então $\overline{x}$ é solução global de (CNP).

Prova:

De (2) e de (3) segue que

$$\overline{\lambda}_i(t) g_i(t, x(t)) \leq 0 = \overline{\lambda}_i(t) g_i(t, \overline{x}(t)) \text{ q.t.p. } [0,T],\ i \in I,\ x \text{ factível.}$$

Como cada $g_i(t,\cdot)$ é invexa em $\overline{x}(t)$ e $\overline{\lambda_i}(t) \geq 0$ q.t.p. em $[0,T]$ segue que $\overline{\lambda}_i(t) g_i(t,\cdot)$ também é invexa em $\overline{x}(t)$, com mesma função η, donde obtemos

$$\overline{\lambda}_i(t) g_i^0(t, \overline{x}(t); \eta(x(t), \overline{x}(t))) \leq 0 \text{ q.t.p. em } [0,T],\ i \in I.$$

Mas $h(\cdot) = \eta(x(\cdot), \overline{x}(\cdot)) \in L_\infty^n[0,T]$, donde segue de (1) que

$$0 \leq \int_0^T [f^0(t, \overline{x}(t); \eta(x(t), \overline{x}(t))) + \sum_{i=1}^m \overline{\lambda}_i(t) g_i^0(t, \overline{x}(t); \eta(x(t), \overline{x}(t)))] dt.$$

Portanto, das duas últimas desigualdades segue que

$$\int_0^T f^0(t, \overline{x}(t); \eta(x(t), \overline{x}(t))) dt \geq 0$$

Como f é invexa em $\overline{x}$ e da última desigualdade, segue que

$$\phi(\overline{x}) \leq \phi(x)$$

para x factível e arbitrário, portanto, $\overline{x}$ é solução global de (CNP). ■

Notemos que, em geral, as condições dos Teoremas 6.2 e 6.3 não podem ser expressas em termos do gradiente generalizado da função Lagrangeana associada a (CNP), ou seja, não garantem que $0 \in \partial_x L(\overline{x}, \overline{\lambda_0}, \overline{\lambda})$.

Entretanto, se tivermos f, g_i regulares em $\overline{x}$, podemos rescrever a candição (1) do Teorema 6.2 como

$$0 \in \partial_x L(\overline{x}, \overline{\lambda_0}, \overline{\lambda}).$$

Dos Teoremas 6.2 e 6.3 e da observação anterior decorrem os seguintes resultados:

Corolário 6.4 *Em (CNP), suponha* $\overline{x} \in \mathbf{F}$ *, que* $f(t, \cdot)$ *é invexa em* $\overline{x}(t)$ *(com respeito a* V*) em* $[0, T]$ *e que para cada* $i \in I$, $g_i(t, \cdot)$ *é estritamente invexa em* $\overline{x}(t)$ *(com respeito a* V*) em* $[0, T]$*, com mesma* $\eta(x(t), \overline{x}(t))$ *para todas as funções. Suponha, ainda, que existam multiplicadores* $\overline{\lambda_0} \in \mathbb{R}$, $\overline{\lambda} = (\overline{\lambda}_1, ..., \overline{\lambda}_m) \in L^m_\infty[0, T]$ *tais que*

$$\begin{gathered} 0 \in \partial_x L(\overline{x}, \overline{\lambda_0}, \overline{\lambda}) \\ \overline{\lambda_0} \geq 0,\ \overline{\lambda}(t) \geq 0 \ q.t.p. \ em \ [0, T] \\ (\overline{\lambda_0}, \overline{\lambda}(t)) \neq 0 \ q.t.p. \ em \ [0, T] \\ \overline{\lambda}_i(t) g_i(t, \overline{x}(t)) = 0 \ q.t.p. \ em \ [0, T] \ para \ todo \ i \in I \end{gathered}$$

Então, $\overline{x}$ *é solução global de (CNP).*

Corolário 6.5 *Em (CNP), suponha* $\overline{x}$ *factível , que* $f(t, \cdot)$ *e* $g_i(t, \cdot)$ *são invexas em* $\overline{x}(t)$ *(com respeito a* V*) em* $[0, T]$ *e que para cada* $i \in I$*, com mesma* $\eta(x(t), \overline{x}(t))$ *para todas as funções. Suponha, ainda, que existam multiplicadores* $\overline{\lambda} = (\overline{\lambda}_1, ..., \overline{\lambda}_m) \in L^m_\infty[0, T]$ *tais que*

$$\begin{gathered} 0 \in \partial_x L(\overline{x}, 1, \overline{\lambda}) \\ \overline{\lambda_0} \geq 0,\ \overline{\lambda}(t) \geq 0 \ q.t.p. \ em \ [0, T] \\ \overline{\lambda}(t) \neq 0 \ q.t.p. \ em \ [0, T] \\ \overline{\lambda}_i(t) g_i(t, \overline{x}(t)) = 0 \ q.t.p. \ em \ [0, T] \ para \ todo \ i \in I \end{gathered}$$

Então, $\overline{x}$ *é solução global de (CNP).*

Capítulo 7

Comentários finais

Eram dois os objetivos do nosso trabalho: primeiro, conhecer o método de deformação e segundo, aplicá-lo ao problema de tempo contínuo, visando a encontrar novas condições suficientes de otimalidade para o problema (CNP).

O primeiro foi cumprido de forma satisfatória; entretanto, o mesmo não se deu con o segundo, pelas razões que explicaremos a seguir.

Para o caso finito-dimensional, como vimos no Capítulo 4, o procedimento adotado por Bobylev foi o seguinte:

- Estudar uma classe de deformação de problemas que preserva a propriedade de um extremo (isto é, um ponto que satisfaz as condições necessárias Fritz John) ser mínimo local (neste caso, as chamadas deformações não singulares);
- Transformar o problema em estudo num de mais simples investigação, via deformação: especificamente, no caso finito- dimensional, Bobylev deforma o problema geral num problema linear.

Nossa tentativa foi a de " imitar " este procedimento em nossa análise do problema (CNP). Portanto, iniciamos um trabalho de busca de deformações de problemas infinito-dimensionais. O melhor que encontramos foi o trabalho de Skalyga [23], cujos principais resultados comentamos a seguir.

O autor considera uma classe de deformação de funcionais definidos em espaços de Hilbert:

Definição 7.1 *Seja $(H, < \cdot, \cdot >)$ um espaço de Hilbert. O funcional $f : H \to \mathbb{R}$ é* ***H- regular*** *se $\nabla f(x)$ satisfaz uma condição Lipschitz em cada bola e também a propriedade (S)*

(S) *Se $(x_n) \subset H$, $x_n \xrightarrow{w} x_0$ e $\limsup_{n\to\infty} < \nabla f(x_n), x_n - x_0 > \leq 0$ então $||x_n - x_0|| \to 0$.*

Skalyga considera a seguinte família parametrizada de problemas:

$$(P(\lambda)) \begin{cases} \text{minimizar } f(x;\lambda) \\ \text{sujeito a: } g_i(x;\lambda) \leq 0, \ i = 1, ..., N \\ x \in H, \lambda \in [0,1] \end{cases}$$

e, para cada $\lambda \in [0,1]$, define extremo da seguinte maneira:

Definição 7.2 *Para cada $\lambda \in [0,1]$, $x_0(\lambda) \in H$ é um* ***extremo*** *de $(P(\lambda))$ se existem $\mu_i(\lambda) \geq 0$, $i = 0, 1, ..., N$ não todos nulos tais que*

- $\nabla_x \mathcal{L}(x_0(\lambda), \mu(\lambda); \lambda) = 0$

  (onde $\mathcal{L}(x_0(\lambda), \mu(\lambda), \lambda) = \mu_0(\lambda) f(x_0(\lambda); \lambda) + \sum_{i=0}^{N} \mu_i(\lambda) g_i(x_0(\lambda), \lambda))$;

- $g_i(x_0(\lambda), \lambda) \leq 0$, $i = 1, ..., N$;

- $\mu_i(\lambda) g_i(x_0(\lambda), \lambda) = 0$, $i = 1, ..., N$.

As deformações que preservam a minimalidade de um extremo são as deformações não singulares, que definimos na seqüência:

Definição 7.3 *A família de problemas $(P(\lambda))$ é uma* ***deformação não singular*** *se:*

- *Para cada $\lambda \in [0,1]$, o funcional $f(x;\lambda)$ é H- regular;*

- *Para cada λ, os funcionais $g_i(x;\lambda)$ são continuamente diferenciáveis em x e $\nabla_x g_i(x;\lambda)$ satisfaz uma condição de Lipschitz em cada bola $T \subset H$, $i = 1, ..., N$;*

- *Os funcionais $f(x;\lambda)$, $g_i(x;\lambda)$ e seus gradientes $\nabla_x f(x;\lambda)$, $\nabla_x g_i(x;\lambda)$ são contínuas em λ, uniformemente com respeito a x em cada bola $T \subset H$, para $i = 1, ..., N$;*

- *Para cada λ, o problema (P(λ)) tem um extremo $x_0(\lambda)$, uniformemente isolado com respeito a λ em H e continuamente dependente de λ;*

- *Para cada λ, $\nabla_x g_i(x_0(\lambda); \lambda)$ $(i = 1, ..., N)$ são linearmente independentes;*

- *Para cada λ o conjunto $\mathcal{D}_\lambda = \{x \in H : g_i(x; \lambda) \leq 0,\ i = 1, ..., N\}$ é convexo.*

Em [23], Skalyga demonstra o seguinte teorema:

Teorema 7.4 *Se (P(λ)) como definida acima é uma deformação não singular e se $x_0(0)$ é mínimo local de (P(0)), então $x_0(1)$ é mínimo lacal de (P(1)).*

Notamos que para o problema (CNP), a noção de extremo- dadas pelas condições necessárias de Fritz John, vistas no Capítulo 5, é bastante diversa da definida por Skalyga- não soubemos contornar esta situação.

Outro fato, ainda mais delicado é que o funcionais do problema (CNP) são definidos num espaço de Banach (e não de Hilbert) e, novamente, não soubemos resolver isto.

Portanto, com o material teórico de que dispúnhamos, não foi possível aplicar (pelo menos diretamente) o método de deformação ao problema (CNP).

Referências

[1] J. P. Aubin e A. Cellina (1984), *"Differential Inclusions "*, Springer- Verlag.

[2] J. P. Aubin e H. Frankowska (1990), *"Set Valued Analysis "*, Birkhauser.

[3] R. Bellman (1953), *"Bottleneck problems and dynamic programming "*, Proc. Nat. Acad. Sci., 39, 947-951

[4] A. J. V. Brandão, M. A. Rojas-Medar e G. N. Silva, (1999), *"Nonsmooth continuous time optimization problems: necesary conditions "* (submetido à publicação)

[5] A. J. V. Brandão (1998), *"Sobre algumas contribuições em otimização não diferenciável invexa "* , Tese de Doutorado em Matemática Aplicada, IMECC-Unicamp.

[6] N. A. Bobylev (1989), "Deformation metohd of investigation of nonlinear programming problems I ", Automation and Remote Control, p. 917-924, v. 50, no. 7, part I.

[7] N. A. Bobylev (1989), "Deformation metohd of investigation of nonlinear programming problems II ", Automation and Remote Control, p. 1018-1026, v. 50, no. 8, part I.

[8] N. A. Bobylev, S. K. Korovin e V. I. Skalyga (1996), *"A homotopic method of studying multivalued problems "*, Automation and Remote Control, vol. 57, no. 10, p. 1513-1521

[9] N. A. Bobylev (1993), *"Homotopic metohods in variational problems "*, Journal of Soviet Mathematics, p. 2659-2712.

[10] N. A. Bobylev (1983) *"Deformation of funcionals having a unique critical point "*, Mat. Zametki, 31, no. 3, 387-398.

-71-

[11] N.A. Bobylev G. V. Kondakov (1991) *"Deformation method for investigation of nonsmooth optimization problems "*, Avtomatika i Telemekhanika, no. 5, 46-57.

[12] H. Brézis (1984), *"Análisis funcional "*, Alianza Universidad de Textos.

[13] F. H. Clarke (1990), *"Optimization and nonsmooth analysis "*, Classics in Applied Mathematics, 5, SIAM.

[14] L. Contesse-Becker (1984), *"Introducción a la optimización con restricciones "*, Universidad de Chile.

[15] B. D. Craven (1995), *"Control and optimization "*, Chapman & Hall.

[16] M. A. Hanson (1981), *"On sufficiency of the Kuhn-Tucker conditions "*, J. Math. Anal. Appl., 80, 545-550.

[17] J. Jahn (1996), *"Introduction to the theory of nonlinear optimization "*, Springer.

[18] O. L. Mangasarian (1969), *"Nonlinear programming "*, Mc Graw-Hill.

[19] M. Rojas-Medar, A. J. V. Brandão e G. N. Silva (1998) *"Nonsmooth continuous time optimization problems: sufficient conditions "*, J. Math. Anal. Appl., 227, 305-318.

[20] M. Rojas-Medar, A. J. V. Brandão e G. N. Silva (1999), *"Introdução a programação matemática "* (notas de um minicurso), IBILCE-Unesp.

[21] H. L. Royden (1968), *"Real Analysis "*, Macmillam.

[22] V. I. Skalyga (1991), *"Deformation Method for investigation of constrained minimization of performance functionals of systems with infinitely many degrees of freedom "*, Automation and remote Control, 52, 781-789.

[23] V. I. Skalyga (1995), *"On deformations of nonsmooth optimization problems having an isolated extremal "*, Izvestiya Math., vol. 45, no. 1.

[24] G. J. Zalmai (1991), *"Optimality conditions and Lagrangian duality in continuous time nonlinear programming "*, J. Math. Anal. Appl., 109, 426-452.

[25] G. J. Zalmai (1985), *"A continuous time generalization of Gordan's Transposition Theorem "*, J. Math. Anal. Appl., 110, 130-140.

[26] G. J. Zalmai (1985), *"The Fritz John Conditions in continuous time nonlinear programming"*, J. Math. Anal. Appl., 110, 503-518.

[27] G. J. Zalmai (1985), *"Sufficiency optimality conditions in continuous time nonlinear programming "*, J. Math. Anal. Appl., 110, 130-147.